1971

Convergence of the difference equation for the error covariance matrix arising in Kalman filter theory

Dennis James Duven
Iowa State University

71-26,851

DUVEN, Dennis James, 1940-
CONVERGENCE OF THE DIFFERENCE EQUATION FOR THE ERROR COVARIANCE MATRIX ARISING IN KALMAN FILTER THEORY.

Iowa State University, Ph.D., 1971
Engineering, electrical

University Microfilms, A XEROX Company, Ann Arbor, Michigan

Convergence of the difference equation for the error covariance matrix arising in Kalman filter theory

by

Dennis James Duven

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Electrical Engineering

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Ames, Iowa

1971

TABLE OF CONTENTS

Page

I. INTRODUCTION

The Kalman filter is a sequential estimator whose input is a set of measurements taken from a random process which can be modeled as a linear state-variable system with white-noise driving functions, and whose output is an estimate of the state of that system at a given instant of time. Kalman [1] was the first to show that such an estimator is also a linear state-variable system whose parameters, or gains, depend upon the solution of a nonlinear matrix difference equation. This equation is called the covariance equation since its solution happens to be the covariance matrix associated with the estimation errors.

The stability of the covariance equation is important because its solution is both an integral part of the Kalman filter and is also indicative of the quality of the estimate. It has been noted in some numerical studies [2, 3] that in some cases the covariance matrix does not remain bounded, and this has generally been attributed to the fact that round-off error is always introduced in any practical implementation of the Kalman filter. However, the manner in which this happens and the circumstances which give rise to it have not been well understood. The purpose of the research reported in this thesis was to investigate the stability properties of this equation and to determine the factors which cause it to be either stable or unstable.

Some theoretical investigations of this problem have been reported by other investigators, and the results of these investigations are reviewed in Chapter II. The additional results obtained by the author are given in Chapter III. The remainder of this chapter is devoted to the precise description of the Kalman filter and to the definition of terms which are used throughout this thesis.

A. The Kalman Filter

The Kalman filter is basically a technique for estimating the value of one or more signals when a set of measurements which are linearly related to those signals, but also contain some error, are known. The technique for doing this at discrete instants of time was published by Kalman [1] in 1960, and the technique for doing this continuously was published by Kalman and Bucy [4] in 1961. The later technique will be described first.[1]

1. The continuous-time Kalman filter

The continuous-time Kalman filter is designed to estimate signals which can be modeled as state variables of a continuous-time dynamic system of the form

[1]State variable notation is used extensively in the description of the Kalman filter, and this notation relies on the use of matrices. For a description of the notational conventions used in this thesis, see Appendix A, and for a description of the state space approach to linear system theory, see Zadeh and Desoer [5].

$$\dot{\underline{x}} = A(t)\underline{x} + \underline{h}(t) \tag{1.1}$$

whose inputs, $\underline{h}(t)$, are white-noise random processes with the statistical properties:

$$E[\underline{h}(t)] = \underline{0} \tag{1.2}$$

$$E[\underline{h}(t)\underline{h}'(\sigma)] = H(t)\delta(t-\sigma), \tag{1.3}$$

where $H(t)$ is a symmetric, nonnegative definite matrix for all $t \geq t_0$ and $\delta(t)$ is the Dirac delta function. The measurements, $\underline{y}(t)$, must also be related to the state variables by an equation of the form

$$\underline{y}(t) = M(t)\underline{x}(t) + \underline{\Delta y}(t) \tag{1.4}$$

where the elements of the vector $\underline{\Delta y}(t)$ are the measurement errors, which must have statistical properties of the form

$$E[\underline{\Delta y}(t)] = \underline{0} \tag{1.5}$$

$$E[\underline{\Delta y}(t)\underline{\Delta y}'(\sigma)] = V(t)\delta(t-\sigma) \tag{1.6}$$

$$E[\underline{\Delta y}(t)\underline{h}'(\sigma)] = 0, \tag{1.7}$$

in which $V(t)$ is also a symmetric, positive definite matrix for all $t \geq 0$.[1]

The Kalman filter is a linear dynamic system described

---

[1]Bryson and Johansen [6] have treated systems in which the inputs and measurement errors are correlated, and in which the $V(t)$ matrix may be positive semidefinite.

by the differential equation

$$\dot{\hat{\underline{x}}} = A(t)\hat{\underline{x}} + K(t)[\underline{y}(t)-M(t)\hat{\underline{x}}], \tag{1.8}$$

where $\hat{\underline{x}}$ is the state of the Kalman filter and is also the optimum linear estimate of $\underline{x}(t)$. The measurements $\underline{y}(t)$ are the inputs to the Kalman filter, the initial state of the filter is

$$\hat{\underline{x}}(t_0) = E[\underline{x}(t_0)], \tag{1.9}$$

and the elements of K(t) are gain variables which are given by the formula

$$K(t) = P(t)M'(t)V^{-1}(t) . \tag{1.10}$$

The n-by-n matrix P(t) is the covariance matrix of the errors in estimation,

$$P(t) = E[(\hat{\underline{x}}-\underline{x})(\hat{\underline{x}}-\underline{x})'], \tag{1.11}$$

and it is the solution of the matrix Riccati equation:

$$\dot{P} = A(t)P + PA'(t) + H(t) - PM'(t)V^{-1}(t)M(t)P . \tag{1.12}$$

A block diagram of the model of the random process and the Kalman filter is shown in Figure 1.1.

The items which must be specified before the Kalman filter can be implemented are: The process coefficients A(t) and M(t), the process covariances H(t) and V(t), the

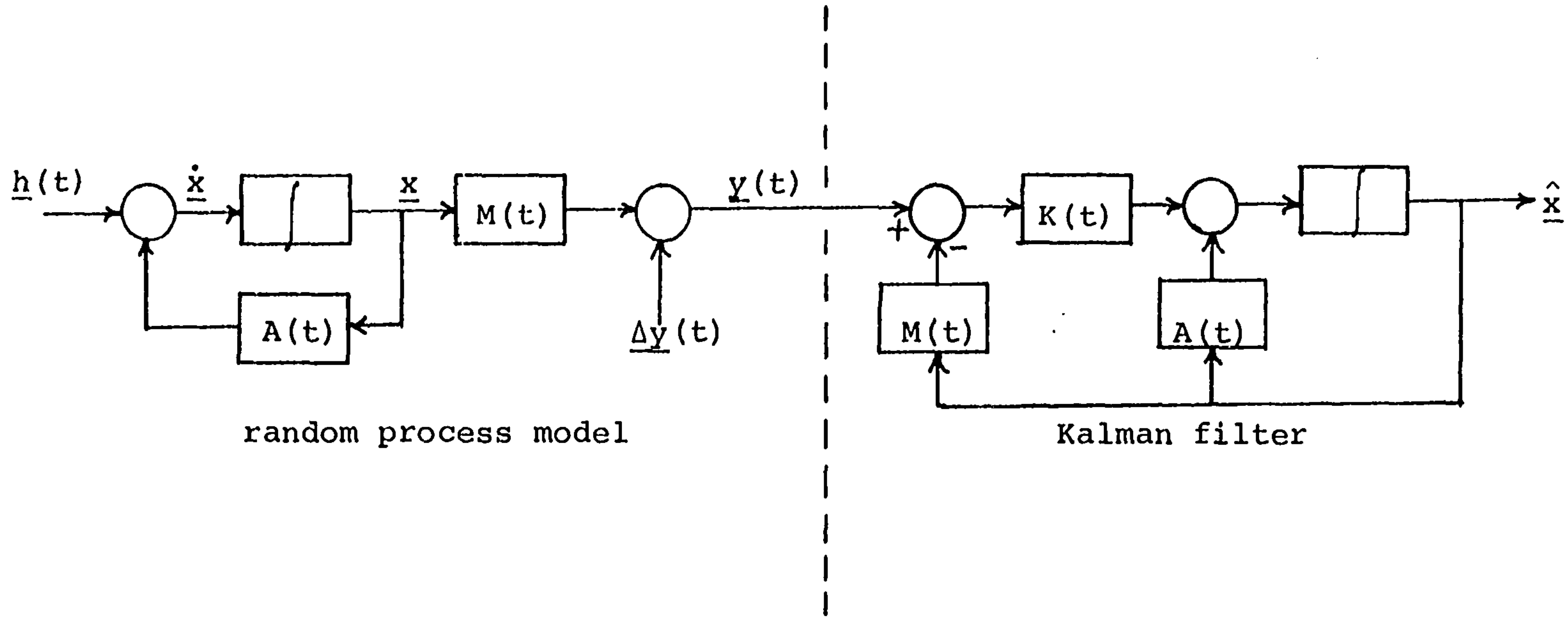


Figure 1.1. Block diagram of the continuous-time random process model and the Kalman filter

initial estimate $\underline{x}(t_0)$, and the covariance of the initial estimate $P(t_0)$.

2. The discrete-time Kalman filter

The discrete-time Kalman filter is designed to estimate $\underline{x}(t)$ only at discrete instants of time $t_1, \ldots, t_k, \ldots$. It is assumed that these signals can be modeled as the state variables of a discrete-time dynamic system of the form[1]

$$\underline{x}_{k+1} = \Phi_k \underline{x}_k + \underline{h}_k \tag{1.13}$$

whose inputs, $\underline{h}_k$, are white noise random sequences with the statistical properties

$$E[\underline{h}_k] = \underline{0} \tag{1.14}$$

$$E[\underline{h}_{k_1} \underline{h}_{k_2}'] = H_{k_1} \delta_{k_1,k_2} \tag{1.15}$$

where H_k is a symmetric, nonnegative definite matrix for all $k \geq 0$, and δ_{k_1,k_2} is the Kronecker delta function. The measurements are assumed to be related to the state variables by the equation

$$\underline{y}_k = M_k \underline{x}_k + \underline{\Delta y}_k \tag{1.16}$$

---

[1]The form $\underline{x}_k$ will be used in preference to the form $\underline{x}(t_k)$ whenever possible. However, when an element of a vector or matrix which depends on t_k is to be denoted, the form $x_i(t_k)$ will be used.

where the elements of $\underline{\Delta y}_k$ are the measurement errors which have the statistical properties

$$E[\underline{\Delta y}_k] = \underline{0} \tag{1.17}$$

$$E[\underline{\Delta y}_{k_1} \underline{\Delta y}_{k_2}'] = V_{k_1} \delta_{k_1,k_2} \tag{1.18}$$

$$E[\underline{\Delta y}_{k_1} \underline{h}_{k_2}'] = 0, \tag{1.19}$$

where V_k is a symmetric, nonnegative definite matrix for all $k \geq 0$.

The discrete-time Kalman filter is the linear dynamic system described by the pair of equations

$$\tilde{\underline{x}}_k = \hat{\underline{x}}_k + K_k[\underline{y}_k - M_k \hat{\underline{x}}_k] \tag{1.20}$$

$$\hat{\underline{x}}_{k+1} = \Phi_k \tilde{\underline{x}}_k , \tag{1.21}$$

where

$\hat{\underline{x}}_k$ = a priori estimate of $\underline{x}_k$, i.e., the estimate based on the measurements $\underline{y}_o \cdots \underline{y}_{k-1}$,

and

$\tilde{\underline{x}}_k$ = a posteriori estimate of $\underline{x}_k$, i.e., the estimate based on the measurements $\underline{y}_0 \cdots \underline{y}_k$.

The gain matrix, in this case, is given by

$$K_k = P_k M_k' (M_k P_k M_k' + V_k)^{-1}, \tag{1.22}$$

where P_k is the a priori covariance matrix

$$P_k \triangleq E[(\hat{\underline{x}}_k-\underline{x}_k)(\hat{\underline{x}}_k-\underline{x}_k)'], \tag{1.23}$$

which is the solution of the pair of equations

$$Q_k = P_k-P_kM_k'(M_kP_kM_k'+ V_k)^{-1}M_kP_k \tag{1.24}$$

$$P_{k+1} = \Phi_kQ_k\Phi_k' +H_k, \tag{1.25}$$

in which Q_k is the a posteriori covariance matrix

$$Q_k \triangleq E[(\tilde{\underline{x}}_k-x_k)(\tilde{\underline{x}}_k-\underline{x}_k)'] . \tag{1.26}$$

Equations (1.24, 1.25) can also be written as the single difference equation

$$P_{k+1} = \Phi_k[P_k-P_kM_k'(M_kP_kM_k' +V_k)^{-1}M_kP_k]\Phi_k' +H_k, \tag{1.27}$$

which is the covariance equation for the discrete-time Kalman filter. A block diagram of the discrete-time random process model and Kalman filter is shown in Figure 1.2.

Equation (1.24) has several alternative forms which are listed below:

$$Q_k = P_k-K_kM_kP_k \tag{1.28a}$$

$$= P_k-P_kM_k'K_k' \tag{1.28b}$$

$$= (I-K_kM_k)P_k(I-K_kM_k)' +K_kV_kK_k' \tag{1.28c}$$

$$= (P_k^{-1}+M_k'V_k^{-1}M_k)^{-1} \tag{1.28d}$$

$$= P_k(I+M_k'V_k^{-1}M_kP_k)^{-1} \tag{1.28e}$$

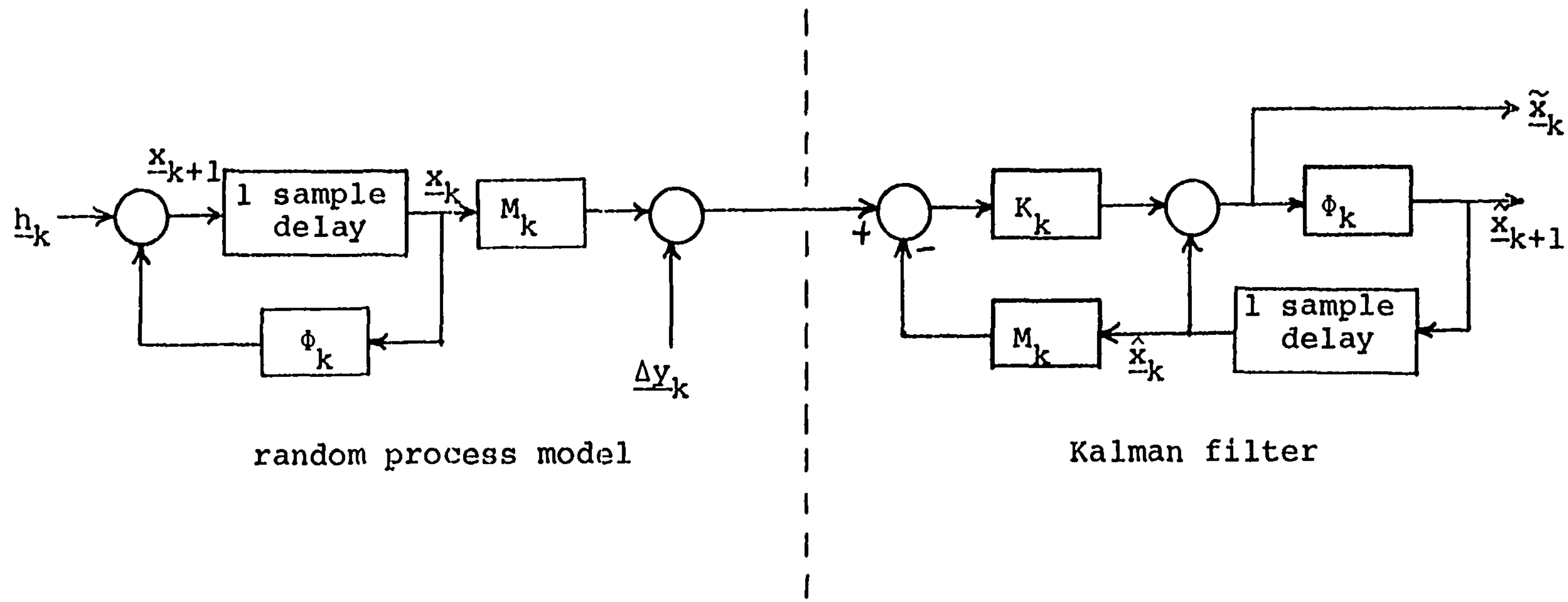


Figure 1.2. Block diagram of the discrete-time random process and Kalman filter

$$= (I+P_kM_k'V_k^{-1}M_k)^{-1}P_k \; . \tag{1.28f}$$

The first two forms are a direct result of (1.22). Form (1.28c) is proved as follows:

$$\begin{aligned}
&(I - K_kM_k)P_k(I - K_kM_k)' + K_kV_kK_k' \\
&= P_k - K_kM_kP_k - P_kM_k'K_k' + K_k(M_kP_kM_k' + V_k)K_k' \\
&= P_k - K_kM_kP_k - P_kM_k'K_k' + P_kM_k'K_k' \\
&= Q_k \; .
\end{aligned} \tag{1.29}$$

Form (1.28d) is called the matrix inversion lemma, and it is proved in [7]. The last two forms are corollaries of the matrix inversion lemma, and they are proved in Appendix B. Obviously, for each form of Q_k, there is a corresponding form of (1.27) which is obtained by replacing the expression within the brackets by one of the right sides of (1.28).

One of the most frequent applications of the discrete-time Kalman filter is to estimate the state of a continuous-time random process, as described by (1.1 - 1.3), when the measurements are made at discrete time instants t_0, t_1, etc. Since the state of a continuous-time system is given by the equation

$$\underline{x}(t) = \Phi(t,t_k)\underline{x}(t_k) + \int_{t_k}^{t} \Phi(t,\tau)\underline{h}(\tau)d\tau \tag{1.30}$$

where $\Phi(t,\tau)$ is the transition matrix from time τ to time t, it is possible to use (1.13) with

$$\Phi_k = \Phi(t_{k+1}, t_k) \tag{1.31}$$

and

$$\underline{h}_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)\underline{h}(\tau)d\tau \tag{1.32}$$

to model the state vector at time t_{k+1}. The Φ_k matrix is nonsingular since it is the transition matrix of a continuous-time system (Zadeh and Desoer [5], p. 340). Also the discrete-time input covariance matrix, H_k, is related to the continuous-time input covariance matrix, $H(t)$, by the equation

$$\begin{aligned} H_k &= E[\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)\underline{h}(\tau)\underline{h}'(\sigma)\Phi'(t_{k+1},\sigma)d\sigma d\tau] \\ &= \int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)H(\tau)\delta(\sigma-\tau)\Phi'(t_{k+1},\sigma)d\sigma d\tau \\ &= \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)H(\tau)\Phi'(t_{k+1},\tau)d\tau \ . \end{aligned} \tag{1.33}$$

B. Controllability and Observability

The concepts of controllability and observability were introduced by Kalman [8, 9] in 1960 and are important in the theory of the Kalman filter.

A continuous-time dynamic system of the form

$$\dot{\underline{x}} = A(t)\underline{x} + G(t)\underline{u}(t) \tag{1.34}$$

is defined to be controllable at time t_0 if for every initial state, $\underline{x}(t_0)$, there exists an input which drives the state to $\underline{0}$ at some finite time $t_1 > t_0$. This is true if and only if the matrix

$$W_a(t_1,t_0) = \int_{t_0}^{t_1} \Phi(t_0,t)G(t)G'(t)\Phi'(t_0,t)\,dt \tag{1.35}$$

is positive definite for some finite $t_1 > t_0$ (Zadeh and Desoer [5], pp. 512-514). The controllability of a time-invariant system of the form

$$\dot{\underline{x}} = A\underline{x} + G\underline{u}(t) \tag{1.36}$$

is independent of the initial time t_0, which implies that it is permissible to call the system controllable rather than controllable at time t_0. Such a system can be tested for controllability by determining if the rank of the matrix

$$Q_c = [G, AG, A^2G, \ldots, A^{n-1}G] \tag{1.37}$$

is equal to n, the dimension of the state space. If the rank of Q_c is less than n, then the system is only partially controllable since only those initial states in the column space of Q_c can be brought to zero in a finite time-interval (Zadeh and Desoer [5], pp. 498-501).

A system which is described by (1.34) and the output equation

$$\underline{y}(t) = M(t)\underline{x}(t) \tag{1.38}$$

is defined to be <u>observable at time t_0</u> if every initial state, $\underline{x}(t_0)$, can be determined from the zero input response over a finite time interval. This is true if and only if the matrix

$$W_b(t_1,t_0) = \int_{t_0}^{t_1} \Phi'(t,t_0)M'(t)M(t)\Phi(t,t_0)dt \tag{1.39}$$

is positive definite for some finite time $t_1 > t_0$. The observability of a time-invariant system described by (1.36) and the output equation

$$\underline{y}(t) = M\underline{x}(t) \tag{1.40}$$

is also independent of t_0, and it can be determined by finding the rank of the matrix

$$P_c = [M',A'M',(A')^2M',\ldots,(A')^{n-1}M']. \tag{1.41}$$

If the rank of P_c is n, the system is <u>observable</u>; if it is less than n, then only the orthogonal projection of $\underline{x}(t_0)$ onto the column space of P_c can be determined from the observed output.

A discrete-time system of the form

$$\underline{x}_{k+1} = \Phi_k\underline{x}_k + G_k\underline{u}_k \tag{1.42}$$

$$\underline{y}_k = M_k \underline{x}_k \tag{1.43}$$

is <u>controllable at time t_0</u> if the matrix

$$W_c(t_k) = \sum_{i=0}^{k-1} \Phi^{-1}(t_i,t_0)G_iG_i'[\Phi^{-1}(t_i,t_0)]' \tag{1.44}$$

is positive definite for some finite $k>0$, where

$$\Phi(t_m,t_j) = \Phi_{m-1}\Phi_{m-2}\cdots\Phi_j \ . \tag{1.45}$$

A time-invariant discrete-time system is <u>controllable</u> if and only if the rank of the matrix

$$Q_d = [G,\Phi G,\Phi^2 G,\ldots,\Phi^{n-1}G] \tag{1.46}$$

is n, and if the rank of Q_d is less than n, then the space of controllable states is the column space of Q_d.

The system is <u>observable at time t_0</u> if the matrix

$$W_d(t_k) = \sum_{i=0}^{k-1} \Phi'(t_i,t_0)M_i'M_i\Phi(t_i,t_0) \tag{1.47}$$

is positive definite for some finite $k>0$, and a time-invariant system is <u>observable</u> if and only if the rank of the matrix

$$P_d = [M',\Phi'M',(\Phi')^2M',\ldots,(\Phi')^{n-1}M'] \tag{1.48}$$

is n. If the rank of P_d is less than n, then only the orthogonal projection of $\underline{x}_0$ onto the column space of P_d can be determined from the observed output.

A random process such as (1.1) is defined to be

<u>completely driven at time t_1</u> if the covariance of the component of the state vector due to the white noise inputs,

$$R_c(t_1) = \int_0^{t_1} \Phi(t_1,\tau)H(\tau)\Phi'(t_1,\tau)d\tau, \tag{1.49}$$

is positive definite (or equivalently, the range space of $R_c(t_1)$ contains the whole state space). Similarly, the discrete-time random process (1.13) is defined to be completely driven at time t_k if the covariance

$$R_d(t_k) = \sum_{i=0}^{k-1} \Phi(t_k,t_{i+1})H_i\Phi'(t_k,t_{i+1}) \tag{1.50}$$

is positive definite.

C. Modes of Linear, Time-Invariant Systems

A study of the modes of a linear system yields considerable insight into the behavior of such systems. Let T be a matrix that transforms A in (1.36) into its Jordan form, $\Lambda=T^{-1}AT$, and let $\underline{w}(t)$ be defined such that

$$\underline{x}(t) = T\underline{w}(t) . \tag{1.51}$$

Then (1.34) is equivalent to the equation

$$\dot{\underline{w}} = \Lambda\underline{w} + T^{-1}G\underline{u}(t), \tag{1.52}$$

with the initial condition

$$\underline{w}(t_0) = T^{-1}\underline{x}(t_0) . \tag{1.53}$$

If A has m distinct eigenvalues, $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$, each of order p_i, then the general form of the Jordan matrix is

$$\Lambda = \mathrm{diag}(\Lambda_1, \Lambda_2, \ldots, \Lambda_m) \tag{1.54}$$

where each submatrix is a p_i-by-p_i matrix of the form

$$\Lambda_i = \begin{bmatrix} \lambda & \delta & 0 & \ldots & 0 & 0 \\ 0 & \lambda_i & \delta & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_i & \delta \\ 0 & 0 & 0 & \ldots & 0 & \lambda_i \end{bmatrix} \tag{1.55}$$

in which each element labelled δ is either one or zero. Let the vector $\underline{w}(t)$ and the matrix T^{-1} be partitioned as

$$\underline{w}(t) = \begin{bmatrix} \underline{\zeta}_1(t) \\ \vdots \\ \underline{\zeta}_m(t) \end{bmatrix} \qquad T^{-1} = \begin{bmatrix} S_1' \\ \vdots \\ S_m' \end{bmatrix} \tag{1.56,1.57}$$

respectively where $\underline{\zeta}_i(t)$ is a p_i-dimensional vector and S_i' is a p_i-by-n submatrix of T^{-1}. Also let the matrix T be partitioned as

$$T = [T_1, T_2, \ldots, T_m] \tag{1.58}$$

in which each T_i is an n-by-p_i submatrix of T, let τ_i be the subspace spanned by the columns of T_i, and let the n-dimensional vector $\underline{\xi}_i(t)$ be defined such that

$$\underline{\xi}_i(t) = T_i \underline{\zeta}_i(t) \ . \tag{1.59}$$

Then (1.51, 1.56, 1.58, 1.59) imply that every vector in the state space can be written as the sum

$$\underline{x}(t) = \sum_{i=1}^{m} T_i \underline{\zeta}_i(t) = \sum_{i=1}^{m} \underline{\xi}_i(t) \tag{1.60}$$

of m vectors, one from each of the subspaces τ_i.

Since Λ is a block diagonal matrix, (1.52) implies that each function $\underline{\zeta}_i(t)$ satisfies the differential equation

$$\dot{\underline{\zeta}}_i = \Lambda_i \underline{\zeta}_i + S_i' G\underline{u}(t) \ , \tag{1.61}$$

and by premultiplying by T_i and replacing $\underline{\zeta}_i$ by

$$\underline{\zeta}_i = S_i' T_i \underline{\zeta}_i = S_i' \underline{\xi}_i \ , \tag{1.62}$$

it can be seen that each function $\underline{\xi}_i(t)$ satisfies the differential equation

$$\dot{\underline{\xi}}_i = T_i \Lambda_i S_i' \underline{\xi}_i + T_i S_i' G\underline{u}(t), \tag{1.63}$$

in which the initial conditions are

$$\underline{\xi}_i(t_0) = T_i S_i' \underline{x}(t_0) \ . \tag{1.64}$$

The vector functions $\underline{\xi}_i(t)$ are called the <u>modes</u> of the system, and they are of interest because, by (1.63), the motion of each mode in the subspace τ_i is independent of the motion of any other mode. The driving function for the ith mode, $T_iS_i'G\underline{u}(t)$, and the initial value of $\underline{\xi}_i$ are respectively the projection of the vectors $G\underline{u}(t)$ and $\underline{x}(t_0)$ onto the subspace τ_i along all other subspaces.

A mode, $\underline{\xi}_i(t)$, of the random process

$$\dot{\underline{x}} = A\underline{x} + \underline{h}(t) \tag{1.65}$$

$$E[\underline{h}(t)] = \underline{0} \tag{1.66}$$

$$E[\underline{h}(t)\underline{h}'(\sigma)] = H\delta(t-\sigma) \tag{1.67}$$

is defined to be completely driven if the subspace τ_i is contained within the range space of the matrix

$$R_c = \int_{t_0}^{t_1} \Phi(t)H\Phi'(t)\,dt \tag{1.68}$$

where t_1 is any number $> t_0$ and R_c is the covariance of the component of the state vector due to the white noise inputs. It is shown in Appendix C that $\underline{\xi}_i(t)$ is completely driven if and only if no eigenvector of A' corresponding to the eigenvalue λ_i^* is a null vector of H. The mode $\underline{\xi}_i(t)$ is defined to be completely observable if $\underline{0}$ is the only vector common to τ_i and $\eta(P_c')$, and it is shown in Appendix C that this is true if and only if no eigenvector of A corresponding to the eigenvalue λ_i is a null vector of M.

The modes of a discrete-time system of the form

$$\underline{x}_{k+1} = \Phi\underline{x}_k + \underline{h}_k \tag{1.69}$$

are similarly defined. Let T be a matrix that transforms Φ into its Jordan form, $\Lambda = T^{-1}\Phi T$, and let (1.51, 1.53-1.60) apply. Then

$$\underline{w}_{k+1} = \Lambda\underline{w}_k + T^{-1}\underline{h}_k \ , \tag{1.70}$$

$$\underline{\zeta}_i(t_{k+1}) = \Lambda_i\underline{\zeta}_i(t_k) + S_i{}'\underline{h}_k \ , \tag{1.71}$$

and the modes are the vectors $\underline{\xi}_i(t_k)$ which satisfy the equation

$$\underline{\xi}_i(t_{k+1}) = T_i\Lambda_i S_i{}'\underline{\xi}_i(t_k) + T_i S_i{}'\underline{h}_k \tag{1.72}$$

and therefore move independently in the subspaces τ_i. The driving functions and initial values are respectively the projection of $\underline{h}_k$ and $\underline{x}_0$ onto τ_i along all other subspaces. When the inputs are white noise with zero mean and covariance H, then the mode $\underline{\xi}_i(t_k)$ is defined to be completely driven if τ_i is contained within the range space of

$$R_d = \sum_{i=0}^{n-1} \Phi^i H(\Phi')^i, \tag{1.73}$$

the covariance at time t_n of the component of the state vector due to the white noise inputs, and it is shown in Appendix C that this is true if and only if no eigenvector of Φ' corresponding to the eigenvalue λ_i^* is a null vector of H. Similarly, the mode is defined to be completely observable if $\underline{0}$ is the only vector common to both τ_i and $\eta(P_d')$, and it is shown in Appendix C that this is true if and only if no eigenvector of Φ corresponding to the eigenvalue λ_i is a null vector of M.

D. Stability of Dynamic Systems[1]

The state equation for a broad category of continuous-time systems can be written in the form

$$\dot{\underline{x}} = \underline{f}(\underline{x},\underline{u}(t),t) \tag{1.74}$$

where $\underline{x}$ is the state vector, $\underline{u}(t)$ is the input vector, and $\underline{f}$ is a vector valued function of the state and input vectors and time. The system is called stationary if $\underline{f}$ does not depend on its third argument, undriven if $\underline{u}(t) = 0$, and autonomous if it is both stationary and undriven. A state, $\underline{x}_e$, of an undriven dynamic system is called an equilibrium state if $\underline{f}(\underline{x}_e, \underline{0}, t) = \underline{0}$ for all t. A system is linear if $\underline{f}$ is linear in its first and second arguments, and (1.74) can then be written in the form

$$\dot{\underline{x}} = A(t)\underline{x} + G(t)\underline{u}(t) \ . \tag{1.75}$$

When the system is undriven, a unique solution of (1.74) exists for all t if $\underline{f}$ is continuous in its third argument and satisfies the global Lipschitz condition

[1]This section is based largely on two companion papers by Kalman and Bertram [10, 11].

$$||\underline{f}(\underline{x}_1,\underline{0},t) - \underline{f}(\underline{x}_2,\underline{0},t)|| \leq k||\underline{x}_1-\underline{x}_2|| \qquad (1.76)$$

for all t, where $\underline{x}_1$ and $\underline{x}_2$ are any two vectors in the state space and k is a positive constant. The value, at time t, of this solution when $\underline{x}(t_0) = \underline{x}_0$ is denoted by $\underline{\phi}(t;\underline{x}_0,t_0)$, and the set of values of $\underline{\phi}(t;\underline{x}_0,t_0)$ as t varies from t_0 to ∞ is called a trajectory of the system.

Many different kinds of stability have been defined, and a given system may be stable according to many or none of the various definitions. Those that are pertinent to this thesis are defined below:

1. Stability. An equilibrium state, $\underline{x}_e$, of an undriven dynamic system is stable if for every positive number ε, there corresponds a positive number $\delta(\varepsilon,t_0)$ such that if $||\underline{x}_0-\underline{x}_e||<\delta$ then $||\underline{\phi}(t;\underline{x}_0,t_0)-\underline{x}_e||\leq\varepsilon$ for all $t\geq t_0$. The notation $\delta(\varepsilon,t_0)$ is used to indicate that, in general, δ depends on ε and t_0.
2. Asymptotic stability. An equilibrium state of an undriven dynamic system is asymptotically stable if (i) it is stable and (ii) if every trajectory starting sufficiently near $\underline{x}_e$ converges to $\underline{x}_e$ as $t\to\infty$, i.e. there exists a positive number $r(t_0)$ such that if $||\underline{x}_0-\underline{x}_e||\leq r(t_0)$, then for every positive number μ there corresponds a time $T(\mu,\underline{x}_0,t_0)$ such that $||\underline{\phi}(t;\underline{x}_0,t_0)-\underline{x}_e||\leq\mu$ for all

$t \geq t_0 + T$.

3. Uniform stability and uniform asymptotic stability. The definitions for uniform stability and uniform asymptotic stability are the same as the definitions for stability and asymptotic stability except in 1) δ is a function of ε only, (t_0 dependence is dropped), and in 2) r is independent of t_0 and T depends on μ and r only.

4. Uniform asymptotic stability in the large. An equilibrium state of an undriven dynamic system is uniformly asymptotically stable in the large if (i) it is uniformly stable, (ii) it is uniformly bounded, i.e. given any positive number r there exists a positive number B(r) such that if $||\underline{x}_0-\underline{x}_e||\leq r$ then $||\underline{\phi}(t;\underline{x}_0 t_0)-\underline{x}_e||\leq B(r)$ for all $t\geq t_0$, (iii) every trajectory converges to $\underline{x}_e$ as $t\to\infty$ uniformly in t_0 and $||\underline{x}_0||\leq r$, i.e. given any positive numbers r and μ there exists a time $T(\mu,r)$ such that if $||\underline{x}_0-\underline{x}_e||\leq r$ then $||\underline{\phi}(t;\underline{x}_0 t_0)-\underline{x}_e||\leq\mu$ for all $t\geq t_0 + T$.

The stability of a system must usually be determined without knowing the general solution of the differential equation. The second method of Lyapunov is a useful and frequently applied tool for this purpose. Its application to autonomous systems of the form

$$\underline{\dot{x}} = \underline{f}(\underline{x}) \tag{1.77}$$

with an equilibrium state at $\underline{x}_e = \underline{0}$ is described by the following theorem.

Theorem 1.1:

The system (1.77) is asymptotically stable in the large if there exists a scalar function, $V(\underline{x})$, with continuous first partial derivatives such that $V(\underline{0}) = 0$ and (i) $V(\underline{x}) > 0$ for all $\underline{x} \neq 0$, (ii) $\dot{V}(\underline{x}) = \nabla V \cdot \underline{f}(\underline{x}) < 0$ for all $\underline{x} \neq 0$, and (iii) $V(\underline{x}) \to \infty$ as $||\underline{x}|| \to \infty$. If condition (iii) is missing, then (1.77) is locally asymptotically stable.

If the system is linear and time-invariant with

$$\dot{x} = A\underline{x}, \tag{1.78}$$

then its stability depends on the eigenvalues of the A matrix as specified by the following two theorems.

Theorem 1.2:

The system (1.78) is stable if and only if (i) all of the eigenvalues of A have nonpositive real parts and (ii) those eigenvalues that lie on the imaginary axis are simple zeros of the minimal polynomial of A.

Theorem 1.3:

The system (1.78) is asymptotically stable if and only if all of the eigenvalues of A have negative real parts.

The state equation for most discrete-time systems can be written in the form

$$\underline{x}(t_{k+1}) = \underline{f}(\underline{x}(t_k),\underline{u}(t_k),t_k) \tag{1.79}$$

where $\underline{x}(t_k)$, $\underline{u}(t_k)$, and $\underline{f}$ have the same significance as they do for continuous-time systems except they apply only at the discrete instants of time t_k. The system (1.79) is <u>stationary</u> if $\underline{f}$ does not depend on its third argument, <u>undriven</u> if $\underline{u}(t_k) = \underline{0}$, and <u>autonomous</u> if it is both stationary and undriven. A state, $\underline{x}_e$, of an undriven discrete-time system is called an <u>equilibrium</u> state if $\underline{f}(\underline{x}_e,\underline{0},t_k) = \underline{x}_e$ for all t_k. The system is <u>linear</u> if $\underline{f}$ is linear in its first and second arguments, and (1.79) can then be written in the form

$$\underline{x}(t_{k+1}) = \Phi_k\underline{x}(t_k) + G_k\underline{u}(t_k) \ . \tag{1.80}$$

A unique solution of (1.79) exists for all t_k if $\underline{f}$ is continuous in its first and third arguments, and the value of this solution at time t_k, when $\underline{x}(t_0) = \underline{x}_0$ and $\underline{u}(t_0) = \underline{0}$, is denoted by $\underline{\phi}(t_k;\underline{x}_0,t_0)$. Except for the replacement of t by t_k, the definitions of stability of a discrete-time system are identical to those of a continuous-time system.

The application of the second method of Lyapunov to autonomous discrete-time systems of the form

$$\underline{x}(t_{k+1}) = \underline{f}(\underline{x}(t_k)) \tag{1.81}$$

with an equilibrium state at $\underline{x}_e = \underline{0}$ is described by the following theorem.

Theorem 1.4:

The system (1.81) is asymptotically stable in the large if there exists a continuous scalar function, $V(\underline{x})$, such that $V(\underline{0}) = 0$ and (i) $V(\underline{x}) > 0$ for all $\underline{x} \neq 0$, (ii) $\Delta V \triangleq V(\underline{x}_{k+1}) - V(\underline{x}_k) < 0$ for all $\underline{x} \neq \underline{0}$, and (iii) $V(\underline{x}) \to \infty$ as $||\underline{x}|| \to \infty$. If condition (iii) is missing, then (1.81) is locally asymptotically stable.

If the system is linear and time-invariant with

$$\underline{x}(t_{k+1}) = \Phi \underline{x}(t_k), \tag{1.82}$$

then its stability depends on the eigenvalues of the Φ matrix as specified by the following two theorems.

Theorem 1.5:

The system (1.82) is stable if and only if (i) all of the eigenvalues of Φ are contained in the closed disk $|\lambda| \leq 1$ and (ii) those eigenvalues that lie on the unit circle are simple zeros of the minimal polynomial of Φ.

Theorem 1.6:

The system (1.82) is asymptotically stable if and only if all of the eigenvalues of Φ are contained in the open disk $|\lambda| < 1$.

II. STATEMENT OF THE PROBLEM, LITERATURE REVIEW, AND RÉSUMÉ OF RESULTS

A. The Problem

Since the solution of the covariance equation is both an integral part of the implementation of the Kalman filter and a measure of the estimation error, the stability of this equation is very important, particularly in systems that operate for long periods of time. Therefore the following questions are of interest in the theory of the Kalman filter.

1. Does a unique solution exist for all $t \geq t_0$, and if so, under what conditions is it stable?
2. Does the initial value of the covariance matrix have any effect on the solution after the system has been operating for a long period of time?
3. Can small errors in the computation of the solution be permitted?
4. What happens if the assumed model of the random process is not accurate?

Other investigators have obtained fairly complete answers to the first, second, and fourth questions in the case of the continuous-time covariance equation when the random process is completely driven and completely observable and several papers have been written on various aspects of the fourth question in the case of the discrete-time covariance equation. The purpose of the research reported in this thesis was to more adequately answer the first three questions in the case of the discrete-time covariance equation, with particular emphasis on the effect of an incompletely driven or incompletely observable random process.

B. Prior Work on the Continuous-Time Covariance Equation

The continuous-time covariance equation has been studied by Kalman [12], Kalman and Bucy[4], and Potter [13], and the following is known about the properties of this equation:

1. (Potter) If $A(t)$, $H(t)$, and $M'(t)V^{-1}(t)M(t)$ are integrable functions for $t \geq t_0$, then (1.12) has a unique solution which is continuous over the interval $t_0 \leq t < t_0 + \tau$ where $\tau > 0$. Since the right side of (1.12) does not satisfy the global Lipschitz condition (1.76), the solution may escape to infinity at a finite time. However if $P(t_0)$ is symmetric nonnegative definite, then $P(t)$ is finite for all $t \geq t_0$.

2. (Potter) If $P(t_0)$ is symmetric nonnegative (positive) definite, then $P(t)$ is symmetric nonnegative (positive) definite for all $t \geq t_0$.

3. (Kalman) If $P(t_0)$ is symmetric nonnegative definite and the random process is completely driven at time $t_1 > t_0$, then $P(t)$ is positive definite for all $t \geq t_1$.

4. (Potter) If $P_a(t)$ and $P_b(t)$ are two solutions of (1.12) with $P_a(t_0) \geq P_b(t_0) \geq 0$, then $P_a(t) \geq P_b(t)$ for all $t \geq t_0$.

In words, if two solutions are initially ordered such that one solution is "more positive definite" than the other, then this ordering is maintained for all time thereafter. As a corollary to this property, Nishimura [14] has shown that the same ordering applies if the inequalities $H_a(t)>H_b(t)$ and $V_a(t)\geq V_b(t)$ are also allowed.

5. (Kalman) If the random process is uniformly completely driven and uniformly completely observable and if $P_a(t)$ and $P_b(t)$ are any two solutions of the covariance equation with nonnegative definite initial values, then $P_a(t)\rightarrow P_b(t)$ as $t\rightarrow\infty$. A system is defined to be <u>uniformly completely driven</u> and <u>uniformly completely observable</u> if there exist fixed positive constants α, β, γ, δ, σ such that

$$\alpha I \leq \int_{t-\sigma}^{t} \Phi(t,\tau)H(\tau)\Phi'(t,\tau)\,d\tau \leq \beta I \tag{2.1}$$

and

$$\gamma I \leq \int_{t-\sigma}^{t} \Phi'(\tau,t-\sigma)M'(\tau)M(\tau)\Phi(\tau,t-\sigma)\,d\tau \leq \delta I \tag{2.2}$$

for all $t\geq\sigma$.

6. (Kalman and Bucy) Let $P(t;0,t_0)$ denote the value of the solution of the covariance equation at time t which corresponds to the initial condition $P(t_0) = 0$. If $A(t)$, $H(t)$, and $M'(t)V^{-1}(t)M(t)$ exist for all t and if for each t there exists a $t_0<t$ such that $W_b(t,t_0)>0$, then

$$P_e(t) = \lim_{t_0 \to \infty} P(t;0,t_0) \tag{2.3}$$

exists and is also a solution of the covariance equation.

7. (Kalman) If the random process is uniformly completely driven and uniformly completely observable, then every solution of the covariance equation whose initial value is nonnegative definite approaches $P_e(t)$ uniformly as $t \to \infty$. Thus $P_e(t)$ is called the <u>moving equilibrium</u> state of the covariance equation. Also under the same conditions, the optimal filter (1.8) is uniformly asymptotically stable.

Potter [13] has studied the properties of the covariance equation when $A(t)$, $M(t)$, $H(t)$, and $V(t)$ are constant matrices equal to A, M, H, and V respectively. The equilibrium solutions (in general there are more than one) are also constant, and are obtained by algebraically solving for the matrices which satisfy the equation

$$H + AP_e + P_eA' - P_eM'V^{-1}MP_e = 0 \ . \tag{2.4}$$

Potter [15] has shown that if the Jordan form of the 2n-by-2n matrix

$$R_c = \begin{bmatrix} A & H \\ M'V^{-1}M & -A' \end{bmatrix} \tag{2.5}$$

is diagonal, then every solution of (2.4) has the form

$$P_e = F\Gamma^{-1} \tag{2.6}$$

where F and Γ are n-by-n matrices equal to the upper and lower halves of the 2n-by-n matrix

$$T = \begin{bmatrix} F \\ \\ \Gamma \end{bmatrix} \tag{2.7}$$

whose columns are selected from the eigenvectors of the R_c matrix. Also every set of eigenvectors of R_c which results in a nonsingular Γ matrix generates a solution to (2.4) as given by (2.6). The eigenvalues of R_c can be shown to be symmetric about the imaginary axis, i.e., if ω is an eigenvalue of R_c, then $-\omega^*$ is also an eigenvalue. In addition, Potter [15] has shown that the set of eigenvalues $\{\omega_1,...,\omega_n\}$ corresponding to the columns of T are uniquely related to the symmetry and definiteness properties of the P_e matrix generated by T as specified by the following two theorems.

Theorem 2.1:

If no two eigenvalues in the set $\{\omega_1,...,\omega_n\}$ satisfy the equation $\omega_i+\omega_j^*=0$ for $1\leq i,j\leq n$, i.e. if no two eigenvalues are mirror images with respect to the imaginary axis, then P_e is Hermitian.

Theorem 2.2:

(a) If P_e is Hermitian positive definite, then the eigenvalues $\{\omega_1,...,\omega_n\}$ have nonnegative real parts. (b) If the eigenvalues $\{\omega_1,\ldots,\omega_n\}$ have positive real parts and Γ is nonsingular, then P_e is Hermitian nonnegative definite.

In [13] Potter defines a random process to be <u>regular</u> if (a) no eigenvector of A, whose corresponding eigenvalue has a nonnegative real part, is a null vector of $M'V^{-1}M$, and (b) no eigenvector of A , whose corresponding eigenvalue has a nonnegative real part, is a null vector of H. Therefore by Appendix C, the random process is regular if no random walk or unstable mode is either undriven or unobservable. Potter shows in [13] that if the system is regular, then a single, stable, nonnegative definite equilibrium solution exists and any solution whose initial value is nonnegative definite approaches this equilibrium solution exponentially fast.

C. Prior Work on the Discrete-Time Covariance Equation

Considerably less has been written on the stability of the discrete-time covariance equation than has been written on the stability of the continous-time covariance equation. Nishimura [16, 17] has studied the effect of using an incorrect P_0 matrix and has shown that if the optimal, actual, and calculated covariances are initially ordered as $P_0(t_0) \leq P_a(t_0) \leq P_c(t_0)$, then this ordering is maintained for all $t_k > t_0$. Heffes [18] has studied the effect of using incorrect H_k and V_k matrices as well as an incorrect P_0 matrix, and has derived a difference equation for the actual covariance matrix in terms of the actual driving and measurement covariances and the transition matrix of the computed filter. Kalman states without proof that all of the properties of the continuous-time covariance equation which he has derived in [12] also apply to the discrete-time case. This has recently been demonstrated, in part, in a paper by Deyst and Price [19] in which the optimal filter (1.20, 1.21) is shown to be uniformly asymptotically stable in the large if the random process (1.13-1.19) is uniformly completely driven and uniformly completely observable. In a companion paper by Price [20], it is shown that these same conditions imply that the difference equation for the actual error covariance matrix is also uniformly asymptotically stable in the large. Although it is not actually proved, the implication of this paper is that the "divergence problem" of the Kalman filter is due to the divergence of the actual covariance matrix from the computed covariance matrix and that this is likely to occur if the random process is not uniformly completely driven and uniformly completely observable.

D. Résumé of Results

Although a few theorems concerning properties of the time-varying discrete-time covariance equation are proved in the next chapter, the primary contribution of this thesis is a detailed analysis of the stability properties of the time-invariant discrete-time covariance equation, similar to Potter's analysis of the time-invariant continuous-time covariance equation. Techniques are derived for algebraically finding the equilibrium solutions of the covariance equation and for predicting their symmetry, definiteness, and local stability. It is shown that the covariance equation has a stable equilibrium solution if the random process has no random walk or unstable mode which is unobservable, and that the stable equilibrium solution is the only nonnegative definite equilibrium solution if the random process is regular. It is also shown in this case that any solution of the covariance equation, whose initial value is nonnegative definite, converges to the stable equilibrium solution. If the random process has an unstable mode which is undriven but observable, then it is shown that the covariance equation has two nonnegative definite equilibrium solutions and that the initial value of a solution determines which of the two equilibrium solutions is approached. Also since one of these is unstable, round-off error generally causes the solution to depart from the unstable equilibrium and to finally approach the stable equilibrium solution.

A somewhat surprising result of the author's research is the fact that if the random process is regular, then the computed covariance matrix does not have to be strictly nonnegative definite, i.e., it is permissible for one or more of the eigenvalues of the covariance matrix to be less than zero. This can happen as the result of round-off error if the equilibrium solution is singular or nearly singular. It is also shown that Price's analysis of the "divergence problem" is correct except that the uniformly completely driven and uniformly completely observable conditions are stronger than necessary.

III. PROPERTIES OF THE DISCRETE-TIME COVARIANCE EQUATION

The stability of the discrete-time covariance equation and related properties are described in this chapter. The existence and uniqueness of a solution is described in Section A, and the symmetry, definiteness, and ordering of solutions is described in Section B. In Section C, the equilibrium solutions and local and global stability of the time-invariant covariance equation are described. While it is certainly a restriction to limit these considerations to the time-invariant case, since in many practical applications the filter is time-varying, the author believes that the results he has obtained do give considerable insight into the stability properties of the covariance equation.

A. Existence and Uniqueness of a Solution

The existence and uniqueness of a solution to the discrete-time covariance equation are more easily proved than in the continuous-time case. If the sampling time, $\Delta t_k = t_{k+1} - t_k$, has a positive lower bound, then the value of the covariance matrix at any finite time is obtained from a finite number of iterations of the covariance equation. Therefore, a unique solution exists at time t_n if the matrix $M_k P_k M_k' + V_k$ is nonsingular for all $k<n$. By Theorem 3.2 of

the next section, the conditions $P_0 \geq 0$ and $V_k > 0$ for all $k<n$, which are true for most applications of the Kalman filter, are sufficient for this to be true. Furthermore, in [12] Kalman shows that even when $M_k P_k M_k' + V_k$ is singular, the solution can be continued by use of the generalized inverse.[1]

B. Symmetry, Definiteness, and Ordering

Since a matrix must be symmetric nonnegative definite to be a covariance matrix, the following two theorems indicate that the solution of the covariance equation has the properties of a covariance matrix if P_0 is symmetric nonnegative definite.

Theorem 3.1:

If P_0 is symmetric, then P_k is symmetric for all $k > 0$.

Proof:

Suppose P_k is symmetric. Then P_{k+1} is symmetric since V_k and H_k are also symmetric. Thus by setting $k = 0, 1, 2, \ldots$ the symmetry of P_k is implied by the symmetry of P_0 for all $k > 0$.

Theorem 3.2:

If P_0 is nonnegative definite, then P_k is nonnegative definite for all $k > 0$.

---

[1] See Penrose [21, 22] for a discussion of the generalized inverse.

Proof:

The covariance equation can be written in the form

$$P_{k+1} = \Phi_k[(I-K_kM_k)P_k(I-K_kM_k)' + K_kV_kK_k']\Phi_k' + H_k \qquad (3.1)$$

by substituting (1.28c) into (1.25). If P_k is nonnegative definite, then P_{k+1} is nonnegative definite since the middle factor in each term of (3.1) is nonnegative definite. Thus, by setting $k = 0, 1, 2, \ldots$ it can be seen that $P_0 \geq 0$ implies $P_k \geq 0$ for all $k > 0$.

The next two theorems show that the discrete-time covariance equation has definiteness and ordering properties identical to those described in items 3 and 4 in Section II.B for the continuous-time case.

Theorem 3.3:

If the discrete-time random process (1.13) is completely driven at time t_m, if $P_0 \geq 0$, and if $V_k > 0$ for all k, then $P_k > 0$ for all $k \geq m$.

Proof:

Suppose P_m were singular with $\underline{z}$ as a null vector. Then by (3.1)

$$\begin{aligned}\underline{z}'P_m\underline{z} &= \underline{z}'\Phi_{m-1}[(I-K_{m-1}M_{m-1})P_{m-1}(I-K_{m-1}M_{m-1})' \\ &\quad + K_{m-1}V_{m-1}K_{m-1}']\Phi_{m-1}'\,\underline{z} + \underline{z}'H_{m-1}\underline{z} \qquad (3.2) \\ &= 0 \;.\end{aligned}$$

Since $P_{m-1} \geq 0$, $V_{m-1} > 0$, and $H_{m-1} \geq 0$, each term in (3.2) must be zero, so $\underline{z}$ must be a null vector of H_{m-1} and $\Phi'_{m-1}\underline{z}$ must be a null vector of K'_{m-1} and P_{m-1}. The above argument can then be repeated to show that $\Phi'_{m-1}\,\underline{z}$ is a null vector of H_{m-2} and $\Phi'_{m-2}\,\Phi'_{m-1}\underline{z}$ is a null vector of K'_{m-2} and P_{m-2}, which implies that the argument can again be repeated. Thus the assumption that $\underline{z}$ is a null vector of P_m implies that $\Phi'_{k+1}\,\ldots\,\Phi'_{m-1}\underline{z} = \Phi'(t_m,t_{k+1})\underline{z}$ is a null vector of H_k for $k = 0, 1, \ldots, m-1$. But this implies that

$$\underline{z}'R_d(t_m)\underline{z} = \sum_{k=0}^{m-1} \underline{z}'\Phi(t_m,t_{k+1})H_k\Phi'(t_m,t_{k+1})\underline{z} = 0, \qquad (3.3)$$

which is impossible since the hypothesis requires the random process to be completely driven at time t_m. Thus $P_m > 0$, which by (3.1) implies that $P_k > 0$ for all $k \geq m$.

<u>Theorem 3.4</u>:

If $V_k > 0$ and $P_a(t_0) \geq P_b(t_0) \geq 0$, then

$P_a(t_k) \geq P_b(t_k)$ for all $k \geq 0$.

<u>Proof</u>:

Let $P_{a0}(\varepsilon) = P_a(t_0) + \varepsilon I$, $P_{b0}(\varepsilon) = P_b(t_0) + \varepsilon I$,

$$P_{a1}(\varepsilon)=\Phi_0[P_{a0}(\varepsilon)-P_{a0}(\varepsilon)M_0'\,(M_0P_{a0}(\varepsilon)M_0'+V_0)^{-1}M_0P_{a0}(\varepsilon)]\Phi_0'+H_0$$

$$= \Phi_0[P_{a0}^{-1}(\varepsilon)+M_0'\,V_0^{-1}M_0]^{-1}\Phi_0'+H_0\,, \qquad (3.4)$$

and

$$P_{b1}(\varepsilon) = \Phi_0[P_{b0}(\varepsilon)-P_{b0}(\varepsilon)M_0'(M_0P_{b0}(\varepsilon)M_0'+V_0)^{-1}M_0P_{b0}(\varepsilon)]\Phi_0' + H_0$$

$$= \Phi_0[P_{b0}^{-1}(\varepsilon)+M_0'V_0^{-1}M_0]^{-1}\Phi_0' + H_0 \ . \qquad (3.5)$$

If $\varepsilon > 0$, then $P_{a0}(\varepsilon) \geq P_{b0}(\varepsilon) > 0$, which implies (Bellman [23], p. 92) that $0 < P_{a0}^{-1}(\varepsilon) \leq P_{b0}^{-1}(\varepsilon)$, which in turn implies that $P_{a1}(\varepsilon) \geq P_{b1}(\varepsilon)$. Also, since the middle terms in (3.4, 3.5) are continuous at $\varepsilon = 0$, $P_a(t_1) = \lim_{\varepsilon\to 0} P_{a1}(\varepsilon)$, $P_b(t_1) = \lim_{\varepsilon\to 0} P_{b1}(\varepsilon)$, and $P_a(t_1) \geq P_b(t_1) \geq 0$. This argument can then be repeated to show that $P_a(t_k) \geq P_b(t_k)$ for $k = 2,3,$ etc.

C. Properties of the Time-Invariant Covariance Equation

In this section, a technique for determining the equilibrium solutions of the time-invariant covariance equation is derived, the properties of these equilibrium solutions are described, some local and global stability theorems are proved, and the stability of the difference equation for the actual covariance matrix is investigated. The matrices Φ_k, M_k, V_k, and H_k are assumed to be independent of k and are denoted by Φ, M, V, and H respectively. Also, it is assumed that V is positive definite and Φ is nonsingular.

The first two subsections are introductory in nature. The first is an analysis of the scalar covariance equation showing the effect of observability, controllability, and stability of the random process, and the effect of the initial variance. The second is an extension of Potter's [15] procedure for the algebraic solution of quadratic matrix equations, and serves as a theoretical basis for several theorems in the third and fourth subsections. The equilibrium solutions and stability of the multidimensional covariance equation are analyzed in subsections three and four respectively.

1. The scalar case

When P_k, Φ, M, V, and H are scalars -- which will be denoted by p_k, ϕ, m, v, and h respectively -- the stability of the covariance equation can be determined from an examination of the graph of p_{k+1} vs. p_k. In terms of the above scalars, the covariance equation takes the form

$$p_{k+1} = \frac{\phi^2 v\, p_k}{m^2 p_k + v} + h \ . \tag{3.6}$$

Thus, when $m \neq 0$, the graph of this equation is a hyperbola whose major axis is parallel to the line $p_{k+1} = -p_k$ and whose asymptotes are the lines $p_k = -v/m^2$ and $p_{k+1} = h + \phi^2 v/m^2$. On the other hand, when $m = 0$, (3.6) becomes the equation

$$p_{k+1} = \phi^2 p_k + h \tag{3.7}$$

whose graph is linear with slope = ϕ^2.

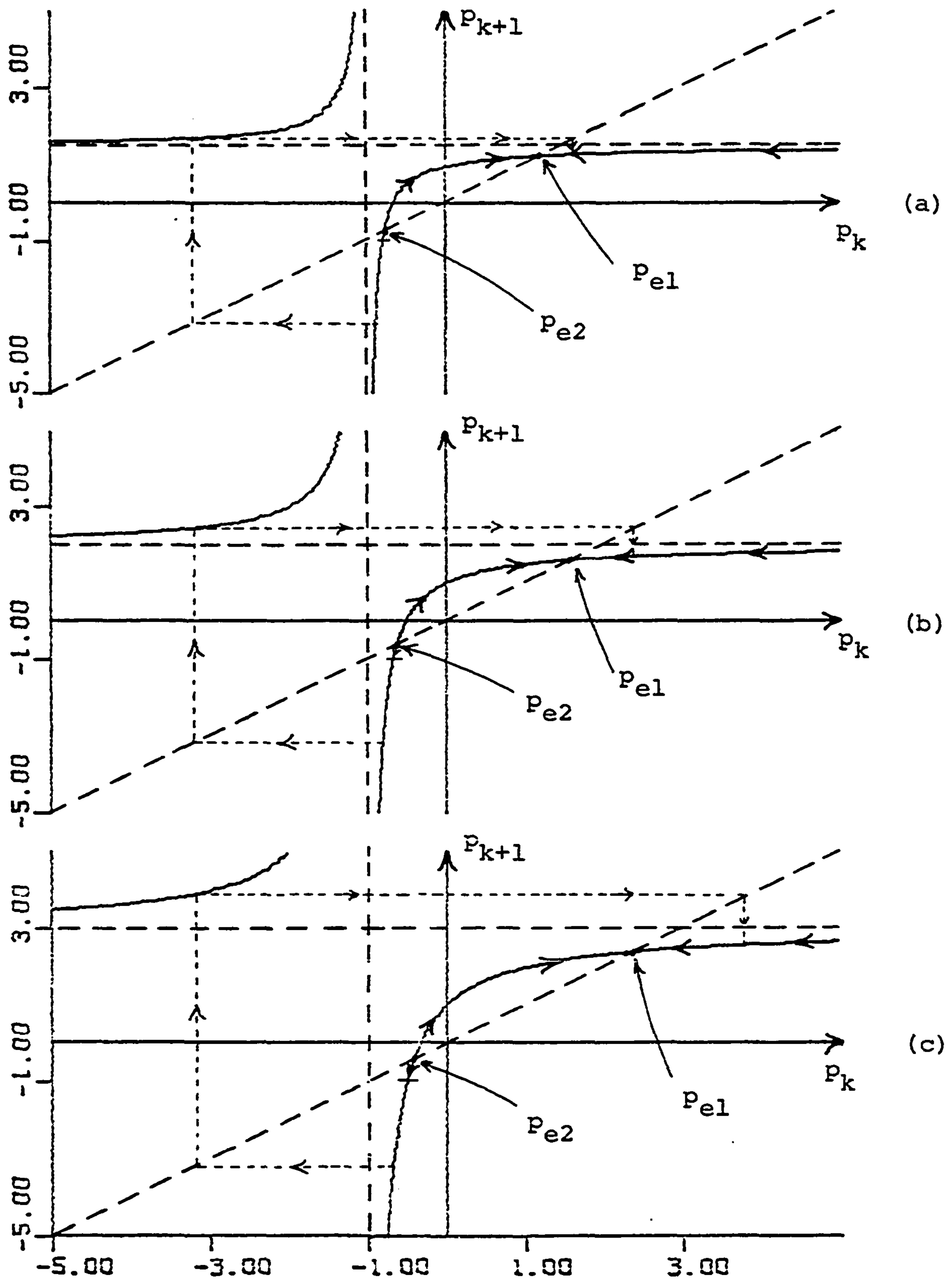
Figures 3.1 (a,b,c) are graphs of (3.6) when the random

Figure 3.1. Graphs of (3.6) when the random process is driven and observable

(a) stable random process

(b) random walk random process

(c) unstable random process



3.00
-1.00
-5.00
p_{k+1}
p_k
p_{e1}
p_{e2}
(a)
(b)
(c)
-5.00
-3.00
-1.00
1.00
3.00

process is driven ($h > 0$) and observable ($m \neq 0$); Figures 3.2 (a,b,c) are graphs of (3.6) when the random process is undriven ($h = 0$) and observable; and Figures 3.3 (a,b,c) are graphs of (3.6) when the random process is driven and unobservable ($m = 0$). Figure (a) in each case corresponds to a stable random process ($\phi^2 < 1$), Figure (b) corresponds to a random walk random process ($\phi^2 = 1$), and Figure (c) corresponds to an unstable random process ($\phi^2 > 1$).

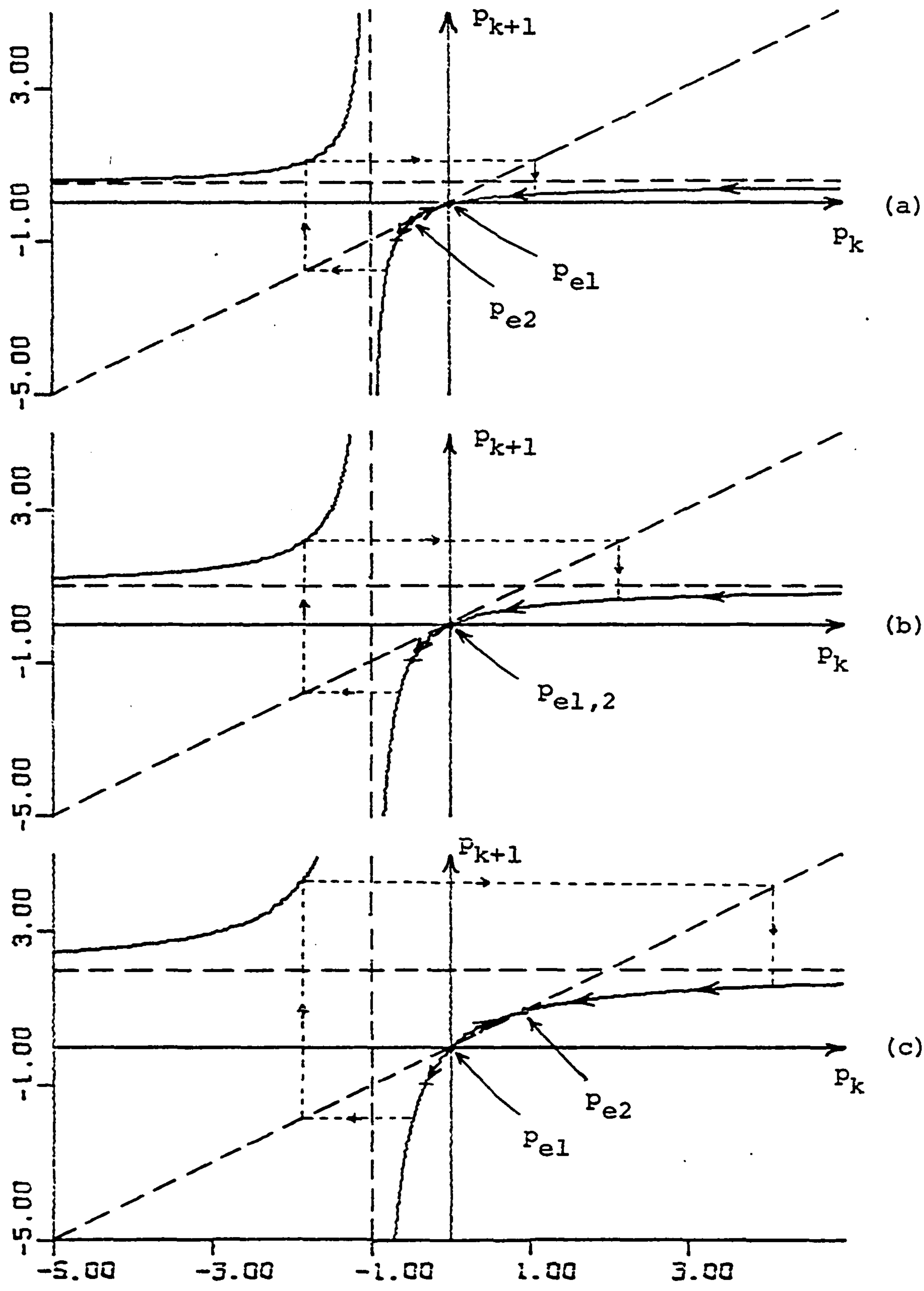
The equilibrium values of p_k are the values at which the p_{k+1} vs. p_k curve intersects the unit diagonal. If p_k is not equal to one of the equilibrium values, then as k increases, the points representing p_k move to the left when the curve is below the unit diagonal and to the right when it is above the unit diagonal. An examination of Figures 3.1 and 3.2 indicates that when the graph is hyperbolic and p_k is sufficiently positive so that the corresponding $p_{k+1} > -v/m^2$, then the points always remain on the lower right section of the curve, so arrows can be placed on this portion of the curve to show the direction of motion. However, if p_k is less than this value and greater than $-v/m^2$, then the next point falls on the upper left section of the hyperbola, and if $p_k < -v/m^2$, then the next point is to the right of the line $p_k = h + \phi^2 v/m^2$. On the other hand, when $m = 0$, there is only one section, so arrows can be placed on the entire curve to show the direction of motion.

Figure 3.2. Graphs of (3.6) when the random process is undriven and observable

(a) stable random process

(b) random walk random process

(c) unstable random process



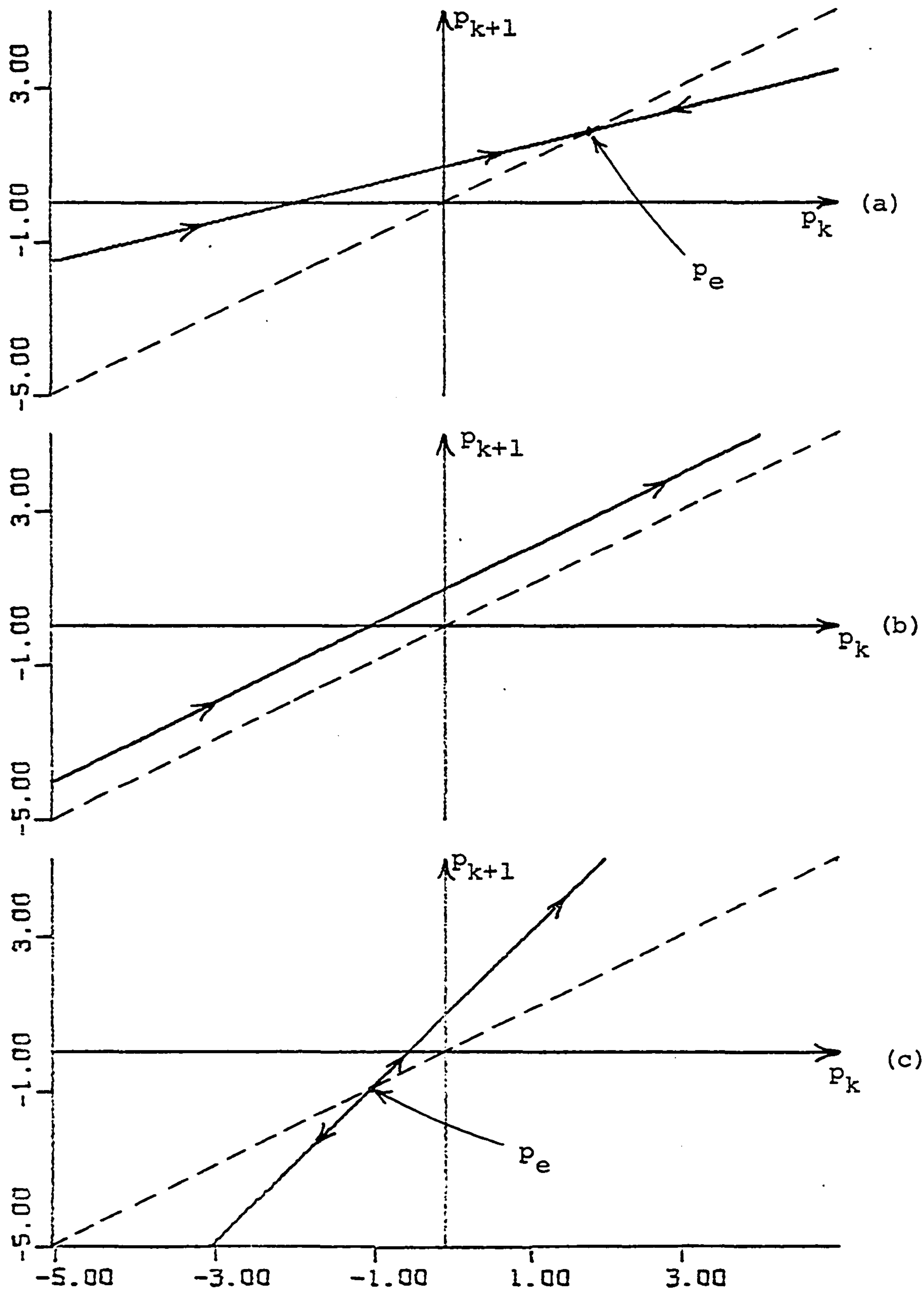
p_{k+1}
3.00
-1.00
-5.00
p_k
p_{e1}
p_{e2}
(a)
p_{k+1}
3.00
-1.00
-5.00
p_k
$p_{e1,2}$
(b)
p_{k+1}
3.00
-1.00
-5.00
p_k
p_{e2}
p_{e1}
(c)
-5.00
-3.00
-1.00
1.00
3.00

Figure 3.3. Graphs of (3.6) when the random process is driven and unobservable

(a) stable random process

(b) random walk random process

(c) unstable random process



p_{k+1}
p_k
(a)
p_e
(b)
(c)
3.00
-1.00
-5.00
1.00
-3.00

The following remarks can now be made about the scalar covariance equation:

(a) When the random process is observable, the equilibrium values are

$$p_{e1,2} = \frac{m^2h + v(\phi^2-1) \pm \sqrt{[m^2h+v(\phi^2-1)]^2+4m^2hv}}{2m^2} \quad . \tag{3.8}$$

Thus when $h > 0$, $p_{e1} > 0$, and $p_{e2} < 0$; and when $h = 0$, then $p_{e1} = 0$ and $p_{e2} = v(\phi^2-1)$ which is positive, zero, and negative for an unstable, random walk, and stable random process respectively.

(b) When the random process is unobservable, there is a single equilibrium at

$$p_e = h(1-\phi^2) \tag{3.9}$$

when the random process is stable or unstable, and no equilibrium when it is a random walk random process.

(c) The scalar covariance equation has a stable, nonnegative equilibrium value unless the random process is unobservable and either random walk or unstable, and except for the undriven unstable case, any solution, whose initial value is nonnegative, converges to the stable equilibrium value. The undriven unstable case has an unstable equilibrium at zero and a stable one which is positive. Therefore, any solution, whose initial value is positive, converges to the stable equilibrium, but the solution whose initial value is exactly zero remains at zero.

(d) Since p_k is a variance, it should never be negative. However, when it is near zero, it is possible for round-off error to make it slightly negative, and if the random process is undriven and either random walk or unstable, the action of the covariance equation is to make the variance even more negative. This eventually results in a catastrophic failure of the Kalman filter. However in all other cases where a stable nonnegative equilibrium solution exists, the action of the covariance equation is to return the variance back toward zero, so no problems result.

2. Solution of the quadratic matrix equation

In the analysis of the multidimensional covariance equation, the solutions to a quadratic matrix equation of the form

$$A + BP + PC - PDP = 0 \tag{3.10}$$

will be needed. A technique for obtaining these solutions is derived in this section. This technique is based on, but represents a significant extension of, Potter's [15] procedure for solving the equation $A + BP + PB^* - PDP = 0$. The following three theorems, which are stated together and then proved provide the theoretical basis for this technique.

Theorem 3.5:

If P satisfies (3.10), then:

(a) There exists a Jordan matrix J, a nonsingular matrix Γ, and a matrix F such that

$$P = F\Gamma^{-1} \tag{3.11}$$

and

$$RT = TJ \tag{3.12}$$

where

$$R = \begin{bmatrix} B & A \\ D & -C \end{bmatrix} \tag{3.13}$$

and

$$T = \begin{bmatrix} F \\ \Gamma \end{bmatrix}; \tag{3.14}$$

(b) There exists a Jordan matrix K, a nonsingular matrix Y, and a matrix Z such that

$$P = -Y^{-1}Z \tag{3.15}$$

and

$$SR = KS \tag{3.16}$$

where

$$S = [Y \quad Z]\ ; \tag{3.17}$$

(c) The matrices F, Γ, Y, and Z satisfy the equation

$$YF + Z\Gamma = 0; \tag{3.18}$$

and

(d) J and K are complements in R, i.e.

$$|\lambda I - R| = |\lambda I - J| \cdot |\lambda I - K| \; . \tag{3.19}$$

<u>Theorem 3.6</u>:

Any solution to (3.12) which has a nonsingular Γ matrix generates a solution to (3.10) through (3.11). Similarly, any solution to (3.16) which has a nonsingular Y matrix generates a solution to (3.10) through (3.15).

<u>Theorem 3.7</u>:

Let $(\lambda-\omega_i)^{r_{ij}}$, $j=1$ to m_i, $i=1$ to σ be the elementary divisors of R, where the integers r_{ij} are ordered such that $r_{ij} \geq r_{i,j+1}$. If T satisfies (3.12) and Γ is nonsingular, then the elementary divisors of J must have the form $(\lambda-\omega_i)^{p_{ij}}$, $j=1$ to n_i, $i=1$ to σ where $p_{ij} \leq r_{ij}, p_{ij} \geq p_{i,j+1}$, and $n_i \leq m_i$ (some of the p_{ij}'s may be zero).

<u>Proof of Theorem 3.5</u>:

(a) Let Γ be a nonsingular matrix which transforms DP-C into its Jordan form

$$J = \Gamma^{-1}(DP-C)\Gamma \; , \tag{3.20}$$

and let

$$F = P\Gamma \ . \tag{3.21}$$

Then by (3.10)

$$A\Gamma + BP\Gamma = P\Gamma\Gamma^{-1}(DP-C)\Gamma \ , \tag{3.22}$$

which by (3.20, 3.21) is equivalent to

$$A\Gamma + BF = FJ \ , \tag{3.23}$$

also by (3.20, 3.21)

$$DF - C\Gamma = \Gamma J \ . \tag{3.24}$$

Thus (3.11) is implied by (3.21) and (3.12) is implied by (3.23, 3.24).

(b) Let Y be a nonsingular matrix which transforms B-PD into its Jordan form

$$K = Y(B-PD)Y^{-1} \ , \tag{3.25}$$

and let

$$Z = -YP \ . \tag{3.26}$$

Then by (3.25, 3.26)

$$YB + ZD = KY \ , \tag{3.27}$$

and by (3.10)

$$YA + YPC = -Y(B - PD)Y^{-1}YP \tag{3.28}$$

which by (3.25, 3.26) is equivalent to

$$YA - ZC = KZ \ . \tag{3.29}$$

Thus (3.15) is implied by (3.26) and (3.16) is implied by (3.27, 3.29).

(c) By (3.11, 3.15)

$$\begin{aligned} YF + Z\Gamma &= Y(F\Gamma^{-1} + Y^{-1}Z)\Gamma \\ &= Y(P-P)\Gamma \\ &= 0 \ . \end{aligned} \tag{3.30}$$

(d) Let E be the matrix

$$E = \begin{bmatrix} F & Y^{-1} \\ \Gamma & 0 \end{bmatrix} \ . \tag{3.31}$$

Then

$$E^{-1} = \begin{bmatrix} 0 & \Gamma^{-1} \\ Y & Z \end{bmatrix} \tag{3.32}$$

and

$$E^{-1}RE = \begin{bmatrix} J & \Gamma^{-1}DY^{-1} \\ 0 & K \end{bmatrix} \ . \tag{3.33}$$

Thus

$$|\lambda I - R| = |\lambda I - E^{-1}RE|$$

$$= |\lambda I - J| \cdot |\lambda I - K| \; . \qquad (3.34)$$

Proof of Theorem 3.6:

If T is a matrix which satisfies (3.12) and has a nonsingular Γ, then (3.23, 3.24) are satisfied, and by postmultiplying both (3.23, 3.24) by Γ^{-1} and premultiplying (3.24) by $P = F\Gamma^{-1}$, the equations

$$BP + A = FJ\Gamma^{-1} \qquad (3.35)$$

$$PDP - PC = FJ\Gamma^{-1} \qquad (3.36)$$

are obtained. Thus by subtracting (3.36) from (3.35), it is seen that P satisfies (3.10). The second part of the theorem is proved in a similar manner.

Proof of Theorem 3.7:

Since Γ is nonsingular, the columns of T are linearly independent. Suppose the elementary divisors of J were not of the stated form. Then either: (a) J has an eigenvalue, μ, not equal to any ω_i, (b) there is an $n_i > m_i$, or (c) there is a $p_{ij} > r_{ij}$. Each of these conclusions leads to a contradiction as shown below:

(a) Let $\underline{x}$ be an eigenvector of J associated with the eigenvalue μ. Then $(RT-TJ)\underline{x} = (R-\mu I)T\underline{x} = \underline{0}$, and since μ is not an eigenvalue of R, $R-\mu I$ is nonsingular and $T\underline{x} = \underline{0}$.

This is a contradiction since the columns of T are linearly independent.

(b) Let the columns of $X = [\underline{x}_1, \ldots, \underline{x}_{n_i}]$ be linearly independent eigenvectors of J associated with the eigenvalue ω_i. Then $(RT-TJ)X = (R-\omega_i I)TX = 0$, which implies that the n_i linearly independent columns of TX are null vectors of $R-\omega_i I$. This is impossible since the dimension of the null space of $R-\omega_i I$ is $m_i < n_i$.

(c) Let E transform R into its Jordan form $\Omega = E^{-1}RE$ and define X such that $T = EX$. Then X satisfies the equation $\Omega X = XJ$, which implies (Gantmacher [24], vol. 1, pp. 215-220) that it has the form

$$X = \mathrm{diag}(X^{(1)}, \ldots, X^{(\sigma)}) \ , \tag{3.37}$$

where $X^{(i)}$ is a $\sum_{j=1}^{m_i} r_{ij}$-by-$\sum_{j=1}^{n_i} p_{ij}$ matrix of the form

$$X^{(i)} = \begin{bmatrix} X_{11}^{(i)} & \ldots & X_{1,n_i}^{(i)} \\ \vdots & & \vdots \\ X_{m_i,1}^{(i)} & & X_{m_i,n_i}^{(i)} \end{bmatrix} . \tag{3.38}$$

The matrices $X_{k\ell}^{(i)}$ are r_{ik}-by-$p_{i\ell}$ regular upper triangular matrices of the form

$$X_{k\ell}^{(i)} = \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \end{bmatrix} \text{ or } X_{k\ell}^{(i)} = \begin{bmatrix} 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{bmatrix}. \qquad (3.39a, 3.39b)$$

if $r_{ik} \geq p_{i\ell}$ if $r_{ik} \leq p_{i\ell}$

Now, for some i, let j be the smallest integer such that $p_{ij} > r_{ij}$, and let $A^{(i)}$ be the submatrix of $X^{(i)}$ whose (k,ℓ) th element is the element in the (1,1) position of $X_{k\ell}^{(i)}$:

$$A^{(i)} = \begin{bmatrix} a_{11} & \cdots & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n_i} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{j-1,1} & \cdots & a_{j-1,j} & a_{j-1,j+1} & \cdots & a_{j-1,n_i} \\ 0 & \cdots & 0 & a_{j,j+1} & \cdots & a_{j,n_i} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{m_i,j+1} & \cdots & a_{m_i,n_i} \end{bmatrix}. \qquad (3.40)$$

The elements in the first j columns on and below the jth row are zero since $r_{ik} < p_{i\ell}$ for all $k \geq j$ and $\ell \leq j$. Thus there exists a linear combination of the first j columns of $A^{(i)}$ which equals $\underline{0}$, and since the elements in $A^{(i)}$ are the only nonzero elements in the corresponding columns of $X^{(i)}$, the same linear combination of the first columns from the first j sets of columns in $X^{(i)}$ is $\underline{0}$. This implies that the columns of X and T are not linear independent, which is a contradiction. Q.E.D.

These theorems, therefore, imply that all of the solutions of the quadratic matrix equation can be generated from the set, $\mathcal{S}$, of all solutions of (3.12) which have nonsingular lower parts. It is possible to rule out all solutions which correspond to J matrices whose elementary divisors do not have the form stated in the conclusion of Theorem 3.7 since these solutions can not have a nonsingular lower part, but all other solutions must be tested for nonsingularity. Also, the following theorem shows that not all of the solutions of (3.12) are needed because many of them generate the same P matrix.

<u>Theorem 3.8</u>:

Let $T_1 = [F_1' \; \Gamma_1']'$ and $T_2 = [F_2' \; \Gamma_2']'$ satisfy the equations $RT_1 = T_1J_1$ and $RT_2 = T_2J_2$ respectively. (a) If T_1 and T_2 are column equivalent, then J_1 and J_2 are similar and $F_1\Gamma_1^{-1} = F_2\Gamma_2^{-1}$. (b) If J_1 and J_2 are not similar, then T_1 and T_2 are not column equivalent and $P_1 = F_1\Gamma_1^{-1}$ is not equal to $P_2 = F_2\Gamma_2^{-1}$.

<u>Proof</u>:

(a) Since T_1 and T_2 are column equivalent, there exists a nonsingular matrix X such that $T_2 = T_1X$. Thus $R(T_1X) = (T_1X)J_2$, $RT_1 = T_1(XJ_2X^{-1})$, and $J_1 = XJ_2X^{-1}$ so J_1 and J_2 are similar. Also $F_2\Gamma_2^{-1} = (F_1X)(\Gamma_1X)^{-1} = F_1XX^{-1}\Gamma_1^{-1} = F_1\Gamma_1^{-1}$.

(b) The first part is simply the contrapositive of the first part of (a). To prove the second part, suppose P_1 were equal to P_2. Then $F_2 = F_1X$, where $X = \Gamma_1^{-1}\Gamma_2$ which is nonsingular, $\Gamma_2 = \Gamma_1X$, and $T_2 = T_1X$. Thus T_1 and T_2 are column equivalent which contradicts the first part. Thus $P_1 \neq P_2$. Q.E.D.

Thus to ensure that every solution in $\mathcal{S}$ generates a unique P matrix, the solutions should be selected such that no two of them are column equivalent. According to the previous theorem, this is no problem for solutions corresponding to nonsimilar J matrices since these solutions are necessarily nonequivalent, but the solutions corresponding to the same J matrix can be column equivalent. In fact when the R matrix is nonderogatory[1] (and in some other cases as well), all of the solutions corresponding to a particular J matrix are column equivalent, so therefore only one of them needs to appear in $\mathcal{S}$. This can be demonstrated by making the change of coordinates T = EX where X is described by (3.37-3.39). If R is nonderogatory then $m_i = 1$, and since $n_i \leq m_i$ and $p_{ij} \leq r_{ij}$, each block in X is a regular upper triangular matrix of the general form

[1]A matrix is defined to be nonderogatory if its minimal and characteristic polynomials are identical. This implies that a nonderogatory matrix has only one elementary divisor per distinguishable eigenvalue, and its Jordan matrix has only one block per distinguishable eigenvalue.

$$X^{(i)} = \begin{bmatrix} a_i & b_i & c_i \\ 0 & a_i & b_i \\ 0 & 0 & a_i \\ 0 & 0 & 0 \end{bmatrix} \tag{3.41}$$

where the parameters a_i, b_i, ... are arbitrary except for the a_i's which must be nonzero since the columns of X are linearly independent. Furthermore, all matrices of the form (3.41) are column equivalent to the cannonical matrix

$$U^{(i)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.42}$$

so therefore it is sufficient to select the matrix

$$T = E\ \mathrm{diag}(U^{(1)},\ldots,U^{(\sigma)}) \tag{3.43}$$

as the only solution in $\mathcal{S}$ corresponding to J. On the other hand if R is derogatory, then the cannonical form of $X^{(i)}$ frequently depends on one or more arbitrary parameters. For example, suppose $m_i = 2$ with $r_{i1} = 2$ and $r_{i2} = 1$, and suppose $n_i = 1$ with $p_{i1} = 2$. Then $X^{(i)}$ would have the general form

$$X^{(i)} = \begin{bmatrix} a_i & b_i \\ 0 & a_i \\ 0 & c_i \end{bmatrix} \tag{3.44}$$

which is column equivalent to the matrix

$$U^{(i)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & \alpha_i \end{bmatrix} \tag{3.45}$$

where α_i is an arbitrary parameter equal to c_i/a_i. Therefore in a case such as this, there are an infinite number of solutions in $\mathcal{L}$ corresponding to the same J matrix, and as a result the quadratic matrix equation has a continuous spectrum of solutions rather than a finite set of isolated solutions.

3. Equilibrium solutions of the covariance equation

The equilibrium solutions of the time-invariant covariance equation are derived in this section and their properties are described. It is shown in part a that the equilibrium matrices satisfy a quadratic matrix equation whose coefficients are functions of Φ, H, M, and V. Thus the equilibrium solutions can be determined by applying the method of the previous section. The resulting R matrix has some special properties which are described in part b. In part c it is shown that the existence, symmetry, and definiteness of an equilibrium solution are uniquely related to the J matrix with which it is associated, and the effects of the stability, observability, and controllability of the random process are described. In part d it is shown that a posteriori equilibrium matrices also satisfy a quadratic matrix equation, and some interesting relationships between the solution of this equation and the quadratic matrix equation for the a priori equilibrium solutions are derived.

a. Derivation of the quadratic matrix equation The equilibrium solutions of the covariance equation are obtained by algebraically solving for the matrices which satisfy the equation

$$P_e = \Phi[P_e - P_e M'(MP_e M' + V)^{-1} MP_e]\Phi' + H \,. \tag{3.46}$$

This equation can be converted to a quadratic matrix equation as follows: First by subtracting H, postmultiplying by $(\Phi')^{-1}$, and applying the identities (1.24, 1.28e), the equation

$$(P_e - H)(\Phi')^{-1} = \Phi P_e (I + M'V^{-1}MP_e)^{-1} \tag{3.47}$$

is obtained. Then by postmultiplying by $I + M'V^{-1}MP_e$ and subtracting ΦP_e, the equation

$$(P_e - H)(\Phi')^{-1}(I + M'V^{-1}MP_e) - \Phi P_e = 0 \tag{3.48}$$

is obtained. And finally by expanding and factoring (3.48) into standard quadratic form, the equation

$$H(\Phi')^{-1} + [H(\Phi')^{-1}M'V^{-1}M + \Phi]P_e$$
$$- P_e(\Phi')^{-1} - P_e[(\Phi')^{-1}M'V^{-1}M]P_e = 0 \tag{3.49}$$

is obtained.

The solutions of (3.49) are obtained by applying the method of Section 2 with

$$A = H(\Phi')^{-1} \tag{3.50a}$$

$$B = H(\Phi')^{-1}M'V^{-1}M + \Phi \tag{3.50b}$$

$$C = -(\Phi')^{-1} \tag{3.50c}$$

$$D = (\Phi')^{-1}M'V^{-1}M \tag{3.50d}$$

and with the R-matrix equal to

$$R_p \triangleq \begin{bmatrix} H(\Phi')^{-1}M'V^{-1}M + \Phi & H(\Phi')^{-1} \\ (\Phi')^{-1}M'V^{-1}M & (\Phi')^{-1} \end{bmatrix}. \tag{3.51}$$

This matrix is equal to the product

$$R_p = \begin{bmatrix} I & H \\ 0 & I \end{bmatrix} \begin{bmatrix} \Phi & 0 \\ 0 & (\Phi')^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ M'V^{-1}M & I \end{bmatrix} \tag{3.52}$$

and since each factor is nonsingular, R_p is nonsingular.

It is possible that extraneous solutions may be introduced when (3.47) is postmultiplied by $I + M'V^{-1}MP_e$. But by (3.50c,d)

$$I + M'V^{-1}MP_e = \Phi'(DP_e - C), \tag{3.53}$$

and by the proof of Theorem 3.5, the matrix $DP_e - C$ is similar to J whose eigenvalues are nonzero since they are also eigenvalues of the nonsingular R_p matrix. Thus $I + M'V^{-1}MP_e$ is nonsingular for every solution of (3.49) and no extraneous solutions are introduced.

b. Characteristics of the R_p matrix   The eigenvalues and eigenvectors of R_p have some special properties which are described in this part. Theorems 3.9 and 3.10 describe a transformation property of the eigenvectors and a symmetry property of the eigenvalues which are true in general, and Theorems 3.11 and 3.12 describe some additional properties which apply if one or more modes of the random process are undriven or unobservable. In the following theorem, let E be a 2n-by-2n matrix whose columns are the eigenvectors of R_p, let Ω be the Jordan form of R_p, and let U and L be n-by-2n matrices equal to the upper and lower halves of E.

Theorem 3.9:

If $\begin{bmatrix} U \\ L \end{bmatrix}$ transforms R_p into its Jordan form, Ω, then $\begin{bmatrix} L \\ -U \end{bmatrix}$ transforms $(R_p^*)^{-1}$ into Ω.

Proof:

By hypothesis

$$\begin{bmatrix} H(\Phi')^{-1}M'V^{-1}M + \Phi & H(\Phi')^{-1} \\ (\Phi')^{-1}M'V^{-1}M & (\Phi')^{-1} \end{bmatrix} \begin{bmatrix} U \\ L \end{bmatrix} = \begin{bmatrix} U \\ L \end{bmatrix} \Omega \,. \qquad (3.54)$$

Therefore by rearranging terms, the equation

$$\begin{bmatrix} (\Phi')^{-1} & -(\Phi')^{-1}M'V^{-1}M \\ -H(\Phi')^{-1} & H(\Phi')^{-1}M'V^{-1}M + \Phi \end{bmatrix} \begin{bmatrix} L \\ -U \end{bmatrix} = \begin{bmatrix} L \\ -U \end{bmatrix} \Omega \qquad (3.55)$$

is obtained, where the matrix on the left is $(R_p^*)^{-1}$. The conclusion follows by observing that the nonsingularity of E implies that $\begin{bmatrix} L \\ -U \end{bmatrix}$ is also nonsingular.

Theorem 3.10:

If ω_i is an eigenvalue of R_p and $\Omega_i = \omega_i I + N_i$ is the associated Jordan block, then $\omega_j = (\omega_i^*)^{-1}$ is also an eigenvalue of R_p and its Jordan block has the form $\Omega_i = \omega_j I + N_j$ with $N_j = N_i$. (N_i and N_j supply the superdiagonal 1's in the Jordan blocks).

Proof:

Let $E_i = \begin{bmatrix} U_i \\ L_i \end{bmatrix}$ be the Jordan vectors associated with ω_i. Then by the proof of Theorem 3.9.

$$(R_p{}^*)^{-1}\begin{bmatrix} L_i \\ -U_i \end{bmatrix} = \begin{bmatrix} L_i \\ -U_i \end{bmatrix}\Omega_i \tag{3.56}$$

and therefore

$$[L_i{}^* \quad -U_i{}^*]R_d = (\Omega_i{}^*)^{-1}[L_i{}^* \quad -U_i{}^*] \ . \tag{3.57}$$

But $(\Omega_i{}^*)^{-1}$ is similar[1] to $(\omega_i{}^*)^{-1}I + N_i$, so there exists a nonsingular matrix X such that

$$X[L_i{}^* \quad -U_i{}^*]R_d = \{(\omega_i{}^*)^{-1}I + N_i\}X[L_i{}^* \quad -U_i{}^*] \ . \tag{3.58}$$

Thus $\omega_j = (\omega_i{}^*)^{-1}$ is also an eigenvalue of R_p and $\omega_j I + N_i$ is a block in Ω associated with ω_j. Furthermore, it is the only block associated with ω_j since if there were another, say $\omega_j I + N_k$, then $(\omega_j{}^*)^{-1}I + N_k$ would be an additional block associated with ω_i, which violates the original assumption on $\omega_i I + N_i$. Thus $N_j = N_i$ and the theorem is proved. It should, however, be mentioned that when ω_i is on the unit circle, then ω_j and ω_i are the same eigenvalue and $\omega_i I + N_i$ is the same diagonal block as $\omega_j I + N_j$. Q.E.D.

This theorem implies that the eigenvalues of R_p have polar symmetry, i.e., each eigenvalue, $\omega = re^{j\theta}$, has an image, $(\omega^*)^{-1} = r^{-1}e^{j\theta}$, with respect to the unit circle. Thus R_p is an exponential function of a matrix whose eigenvalues are images with respect to the imaginary axis, similar

[1]See Appendix B.

to the rectangular symmetry of R_c. Because of this and the form of (3.52), it would seem reasonable for the R_p matrix to be equal to the expression

$$\exp\left(\begin{bmatrix} 0 & H \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} A & 0 \\ 0 & -A^* \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ M\,V^{-1}M & 0 \end{bmatrix}\right)$$

$$= \exp\left(\begin{bmatrix} A & H \\ M\,V^{-1}M & -A^* \end{bmatrix}\right) \tag{3.59}$$

where $A = \ln(\Phi)$, but numerical examples show that this is not true. The reason for this is the fact that $e^{X+Y} \neq e^X e^Y$ unless X and Y commute. Thus R_p is not equal to (3.59) since the matrices in it do not usually commute.

Theorem 3.11:

If a mode, whose eigenvalue is λ, is either undriven or unobservable, then λ and $(\lambda^*)^{-1}$ are eigenvalues of R_p.

Proof:

By Appendix C,

$$\Phi' N_d = N_d \Lambda_d^* \tag{3.60}$$

and

$$HN_d = 0 \ , \tag{3.61}$$

where the columns of N_d are basis vectors of $\eta(R_d)$ and Λ_d is a Jordan matrix whose diagonal elements are equal to the eigenvalues of the undriven modes. Similarly,

$$\Phi N_0 = N_0 \Lambda_0 \tag{3.62}$$

and

$$MN_0 = 0, \tag{3.63}$$

where the columns of N_0 are basis vectors of $\eta(P_d^{\prime})$ and Λ_0 is a Jordan matrix whose diagonal elements are equal to the eigenvalues of the unobservable modes. The equations

$$R_p \begin{bmatrix} 0 \\ \\ N_d \end{bmatrix} = \begin{bmatrix} 0 \\ \\ N_d \end{bmatrix} (\Lambda_d^*)^{-1} \tag{3.64}$$

and

$$[N_d^* \quad 0] R_p = \Lambda_d [N_d^* \quad 0] \tag{3.65}$$

are implied by (3.60), (3.61), and the equations

$$R_p = \begin{bmatrix} N_0 \\ \\ 0 \end{bmatrix} = \begin{bmatrix} N_0 \\ \\ 0 \end{bmatrix} \Lambda_0 \tag{3.66}$$

and

$$[0 \quad N_0^*] R_p = (\Lambda_0^*)^{-1} \; [0 \quad N_0^*] \tag{3.67}$$

are implied by (3.62, 3.63). Therefore if the mode is undriven, λ is a diagonal element of Λ_d and (3.64, 3.65) imply that λ and $(\lambda^*)^{-1}$ are eigenvalues of R_p; if the mode is unobservable, then λ is a diagonal element of Λ_0 and (3.66, 3.67) imply that λ and $(\lambda^*)^{-1}$ are eigenvalues of R_p.

Theorem 3.12:

The R_p matrix has an eigenvalue on the unit circle if and only if the random process has a random walk mode which is either undriven or unobservable.

Proof:

→Let ω be an eigenvalue which is on the unit circle and let $\begin{bmatrix} \underline{f} \\ \underline{\gamma} \end{bmatrix}$ be the corresponding eigenvector. Then $\underline{f}$ and $\underline{\gamma}$ satsify the equations

$$\Phi\underline{f} \qquad = \omega(\underline{f} - H\underline{\gamma}) \tag{3.68}$$

$$M'V^{-1}M\underline{f} = \omega\Phi'\underline{\gamma} - \underline{\gamma} \quad , \tag{3.69}$$

and since $\omega^*\omega = 1$, (3.68) is equivalent to

$$H\underline{\gamma} = \underline{f} - \omega^*\Phi\underline{f} \ . \tag{3.70}$$

Next if (3.69) is premultiplied by $\underline{f}^*$ and (3.70) is premultiplied by $\underline{\gamma}^*$ and transposed, the equations

$$\underline{f}^*M'V^{-1}M\underline{f} = \omega f^*\Phi'\underline{\gamma} - \underline{f}^*\underline{\gamma} \tag{3.71}$$

$$\underline{\gamma}^*H\underline{\gamma} \qquad = \underline{f}^*\underline{\gamma} - \omega\underline{f}^*\Phi'\underline{\gamma} \tag{3.72}$$

are obtained, which imply that

$$\underline{f}^*M'V^{-1}M\underline{f} + \underline{\gamma}^*H\underline{\gamma} = 0 \ . \tag{3.73}$$

Thus since both $M'V^{-1}M$ and H are nonnegative definite, $M\underline{f} = \underline{0}$ and $H\underline{\gamma} = 0$, and (3.69, 3.70) imply the equations

$$\Phi'\underline{\gamma} = \omega^*\underline{\gamma} \tag{3.74}$$

$$\Phi\underline{f} = \omega\underline{f} \tag{3.75}$$

respectively. Therefore, since $\underline{\gamma}$ and $\underline{f}$ can not both be zero, the random process has a random walk mode which is either undriven or unobservable.

← This is a simple application of the previous theorem.

c. Existence, symmetry, and definiteness of the equilibrium solutions The first five theorems in this section show that the existence, symmetry, and definiteness properties of an equilibrium solution are related to the J matrix with which it is associated. Some of these theorems do not apply when the R_p matrix is derogatory. This occurs most commonly when there is a random walk mode which is both undriven and unobservable, but it can also happen when the random process has two modes whose eigenvalues satisfy the equation $\lambda_i^*\lambda_j = 1$ and one of them is undriven and the other is unobservable. In such cases, the existence, symmetry, and definiteness properties must be determined by direct evaluation of the equilibrium solutions. These cases, however, are the exception rather than the rule.

In the first theorem, some conditions which are sufficient to prevent the existence of an equilibrium solution due to a singular Γ matrix are listed. Although no proof is given, these conditions are believed to be necessary as well as sufficient since no exceptions in over 300 numerical solutions have occurred.

Theorem 3.13:

When the R_p matrix is nonderogatory, Γ is singular if (a) there is an eigenvalue in J equal to an eigenvalue of Λ_0, or (b) there is an eigenvalue in $(\Lambda_0^*)^{-1}$ which is not an eigenvalue of J.

Proof:

(a) Let λ_0 be an eigenvalue of both J and Λ_0. Then by (3.66), there is a column of T proportional to $\begin{bmatrix} \underline{n}_0 \\ \underline{0} \end{bmatrix}$, where $\underline{n}_0$ is an element of $\eta(P_d')$, since the subspace of solutions to the equation $R_p \begin{bmatrix} \underline{f} \\ \underline{\gamma} \end{bmatrix} = \lambda_0 \begin{bmatrix} \underline{f} \\ \underline{\gamma} \end{bmatrix}$ is one dimensional when R_p is nonderogatory. Thus Γ is singular since one of its columns is zero.

(b) Let $(\lambda_0^*)^{-1}$ be an eigenvalue of $(\Lambda_0^*)^{-1}$ which is not an eigenvalue of J. Then by (3.67), the equation

$$[\underline{0}^* \quad \underline{n}_0^*] R_p \begin{bmatrix} F \\ \Gamma \end{bmatrix} = [\underline{0}^* \quad \underline{n}_0^*] \begin{bmatrix} F \\ \Gamma \end{bmatrix} J \tag{3.76}$$

is equivalent to the equation

$$\underline{n}_0^* \Gamma (J - (\lambda_0^*)^{-1} I) = \underline{0}^* \ , \tag{3.77}$$

which implies that $\underline{n}_0^* \Gamma = \underline{0}^*$ since $J - (\lambda_0^*)^{-1} I$ is nonsingular if $(\lambda_0^*)^{-1}$ is not an eigenvalue in J. Thus Γ is singular since it has a null row vector. Q.E.D.

The following theorem lists in a similar manner the conditions which result in a singular F matrix.

Theorem 3.14:

Let λ_d be the eigenvalue of an undriven mode. When R_p is nonderogatory, F is singular if (a) $(\lambda_d^*)^{-1}$ is an eigenvalue in J, or (b) λ_d is not an eigenvalue in J. Furthermore if $P_e = F\Gamma^{-1}$ exists, it is singular and $P_e\underline{n}_d = \underline{0}$ in case (a) and $\underline{n}_d^* P_e = \underline{0}^*$ in case (b), where $\underline{n}_d$ is the eigenvector of Φ' in $\eta(R_d)$ corresponding to λ_d^*.

Proof:

(a) Suppose the ith element on the diagonal of J is equal to $(\lambda_d^*)^{-1}$. Then by (3.64) and the assumption that R_p is nonderogatory, the ith column of T is proportional to $\begin{bmatrix}\underline{0}\\ \underline{n}_d\end{bmatrix}$, so F is singular since it has a zero column. Also

$$
\begin{aligned}
P_e\underline{n}_d &= [\underline{f}_1,\ldots,\underline{0},\ldots,\underline{f}_n][\underline{\gamma}_1,\ldots,\underline{n}_d,\ldots,\underline{\gamma}_n]^{-1}\underline{n}_d \\
&= [\underline{f}_1,\ldots,\underline{0},\ldots,\underline{f}_n]\begin{bmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{bmatrix} \qquad (3.78)\\
&= \underline{0}
\end{aligned}
$$

where the elements $\underline{0}$, $\underline{n}_d$, and 1 occupy the ith position in

F, Γ, and $\Gamma^{-1}\underline{n}_d$ respectively.

(b) By (3.65) the equation

$$[\underline{n}_d^* \quad \underline{0}^*]R_p \begin{bmatrix} F \\ \Gamma \end{bmatrix} = [\underline{n}_d^* \quad \underline{0}^*] \begin{bmatrix} F \\ \Gamma \end{bmatrix} J \tag{3.79}$$

is equivalent to the equation

$$\underline{n}_d^* F(J-\lambda_d I) = \underline{0}^* , \tag{3.80}$$

which implies that if λ_d is not an eigenvalue in J, then $\underline{n}_d^* F = \underline{0}^*$ and F is singular. Also if Γ is nonsingular, then $\underline{n}_d^* P_e = \underline{n}_d^* F\Gamma^{-1} = \underline{0}^*$. Q.E.D.

The following two theorems describe how the symmetry of an equilibrium solution is determined by its associated J matrix. In these theorems, ψ_i and ω_i denote the i<u>th</u> distinct eigenvalue of J and Ω respectively and the integers $p_{i1} \geq \cdots \geq p_{i,n_i}$ and $r_{i1} \geq \cdots \geq r_{i,m_i}$ denote the dimensions of the simple Jordan blocks in J and Ω which are associated with i<u>th</u> eigenvalue. Also it is assumed that the eigenvalues in Ω are ordered such that $\omega_i = \psi_i$.

<u>Theorem 3.15</u>:

An equilibrium solution is Hermitian if (a) $\psi_i^* \psi_j \neq 1$ for any two eigenvalues in J (i=j not excluded), or (b) $p_{i1} + p_{j1} \leq r_{i,m_i}$ for every (i, j) pair for which $\psi_i^* \psi_j = 1$.

Proof:

By (3.11, 3.12, 3.14), $P_e = F\Gamma^{-1}$ where

$$R_p \begin{bmatrix} F \\ \Gamma \end{bmatrix} = \begin{bmatrix} F \\ \Gamma \end{bmatrix} J \ . \tag{3.81}$$

Also by (3.54),

$$\begin{bmatrix} F \\ \Gamma \end{bmatrix} = \begin{bmatrix} U \\ L \end{bmatrix} X \tag{3.82}$$

where X is a 2n-by-n matrix which satisfies the equation

$$\Omega X = XJ \ . \tag{3.83}$$

Let $Q = \Gamma^* F$; then Q is Hermitian if and only if

$$\begin{aligned} Q - Q^* &= \Gamma^* F - F^* \Gamma \\ &= X^*(L^* U - U^* L)X \\ &= 0 \ . \end{aligned} \tag{3.84}$$

By premultiplying (3.54) by R_p^{-1} and taking the conjugate transpose of (3.55), it can be seen that

$$L^* U - U^* L = \Omega^*(L^* U - U^* L)\Omega \ , \tag{3.85}$$

and since $\Omega = \mathrm{diag}(\Omega_1, \ldots, \Omega_\sigma)$ where Ω_i is the set of Jordan blocks associated with ω_i, $L^* U - U^* L$ can be partitioned into blocks $W_{\mu\nu}$ which satisfy the equation

$$W_{\mu\nu} = \Omega_\mu{}^* W_{\mu\nu}\Omega_\nu \ . \tag{3.86}$$

When $\omega_\mu{}^*\omega_\nu \neq 1$, the only solution to (3.86) is $W_{\mu\nu} = 0$. Also since $J = \mathrm{diag}(J_1,\ldots,J_s)$, X can be partitioned into blocks X_{ij} which satisfy the equation

$$\Omega_i X_{ij} = X_{ij} J_j \tag{3.87}$$

whose only solution when $\omega_i \neq \psi_j$ is $X_{ij} = 0$. Finally, $Q - Q^*$ can be partitioned into blocks

$$Q_{ij} = \sum_{\mu=1}^{\sigma} \sum_{\nu-1}^{\sigma} X_{\mu i}{}^* W_{\mu\nu} X_{\nu j}$$

$$= X_{ii}{}^* W_{ij} X_{jj} \ , \tag{3.88}$$

which are all zero when $\psi_i{}^*\psi_j \neq 1$. But $Q = \Gamma^* P_e \Gamma$ which is Hermitian if and only if P_e is. Thus under the conditions of part (a), P_e is Hermitian.

Now suppose $\psi_i{}^*\psi_j = 1$. Since Ω_i is the direct sum of the $r_{i\mu}$-dimensional, simple Jordan blocks $\Omega_{i\mu}$, W_{ij} can be partitioned into blocks $W_{\mu\nu}^{(i,j)}$, with dimensions $r_{i\mu}$-by-$r_{j\nu}$, which satisfy the equation

$$W_{\mu\nu}^{(i,j)} = \Omega_{i\mu}{}^* W_{\mu\nu}^{(i,j)} \Omega_{j\nu} \tag{3.89}$$

and consequently have the form

$$W_{\mu\nu}^{(i,j)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & * \\ 0 & * & * \\ * & * & * \end{bmatrix} \quad \text{or} \quad = \begin{bmatrix} 0 & 0 & 0 & * \\ 0 & 0 & * & * \\ 0 & * & * & * \end{bmatrix} \tag{3.90}$$

$$\text{if } r_{i\mu} > r_{j\nu} \qquad\qquad \text{if } r_{i\mu} < r_{j\nu} \ ,$$

where the asterisks indicate nonzero elements. Also since J_i is the direct sum of the $p_{i\nu}$-dimensional, simple Jordan blocks $J_{i\nu}$, X_{ii} (and X_{jj}) can be partitioned into $r_{i\mu}$-by-$p_{i\nu}$ dimensional blocks of the form

$$X_{\mu\nu}^{(i,j)} = \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \end{bmatrix} \quad \text{or} \quad = \begin{bmatrix} 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{bmatrix} \tag{3.91}$$

$$\text{if } r_{i\mu} > p_{i\nu} \qquad\qquad \text{if } r_{i\mu} < p_{i\nu} \ ,$$

where a, b, c, ... represent arbitrary parameters. Thus Q_{ij} is composed of $p_{i\alpha}$-by-$p_{j\beta}$ dimensional blocks of the form

$$Q_{\alpha\beta}^{(i,j)} = \sum_{\mu=1}^{m_i} \sum_{\nu=1}^{m_i} (X_{\mu\alpha}^{(i,i)})^* W_{\mu\nu}^{(i,j)} X_{\nu\beta}^{(j,j)} \ . \tag{3.92}$$

The (μ,ν)th term in (3.92) is zero if

$$\min(p_{i\alpha}, r_{i\mu}) + \min(r_{j\nu}, p_{j\beta}) \leq \max(r_{i\mu}, r_{j\nu}) \ . \tag{3.93}$$

Since $p_{i1} \geq p_{i\alpha}$, $r_{i\mu} \geq r_{im_i}$, $r_{j\nu} \geq r_{jm_j}$,

$p_{j\beta} \leq p_{j1}$, and $p_{i1} + p_{j1} \leq r_{im_i} = r_{jm_j}$, (3.93) is true for every term in (3.92). Thus, every block in Q_{ij} is zero, and if this is true for every (i,j) pair, then $Q - Q^* = 0$ and P_e is Hermitian.

Theorem 3.16:

An equilibrium solution is symmetric if (a) $\psi_i \psi_j \neq 1$ for any two eigenvalues in J, or (b) $p_{i1} + p_{j1} \leq r_{im_i}$ for every (i, j) pair for which $\psi_i \psi_j = 1$.

Proof:

The proof is identical to the previous proof except that all conjugate transposes are replaced by ordinary transposes and all complex conjugates in scalar equations such as $\psi_i^* \psi_j = 1$ are deleted.

An obvious corollary of the preceding two theorems is the fact that P_e is real and symmetric if J satisfies the conditions of both theorems. Although these conditions are sufficient but not necessary, numerical examples indicate that exceptions occur only when R_p is derogatory. The next theorem describes the relationship between the definiteness of an equilibrium solution and its corresponding J matrix.

Theorem 3.17:

When R_p is nonderogatory, a Hermitian equilibrium solution is nonnegative definite if and only if each eigenvalue in J is either (a) greater than one in absolute value, or (b) equal to the conjugate reciprocal of the eigenvalue of an undriven mode.

<u>Proof</u>:

$\leftarrow$ Let the diagonal blocks in J be arranged such that

$$J = \mathrm{diag}(J_1, J_2) \tag{3.94}$$

where the eigenvalues in J_1 are greater than one in absolute value and each eigenvalue in J_2 is an eigenvalue of $(\Lambda_d^*)^{-1}$. Also let T be partitioned such that

$$R_p \begin{bmatrix} F_1 & F_2 \\ \Gamma_1 & \Gamma_2 \end{bmatrix} = \begin{bmatrix} F_1 & F_2 \\ \Gamma_1 & \Gamma_2 \end{bmatrix} \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} \tag{3.95}$$

where $F_2 = 0$ since when R_p is nonderogatory, (3.64) implies that

$$\begin{bmatrix} F_2 \\ \Gamma_2 \end{bmatrix} = \begin{bmatrix} 0 \\ N_d \end{bmatrix} X \tag{3.96}$$

where X satisfies the equation $(\Lambda_d^*)^{-1}X = XJ_2$. Therefore since $Q = \Gamma^* P_e \Gamma$ and P_e is Hermitian, Q has the form

$$Q = \begin{bmatrix} Q_{11} & 0 \\ 0 & 0 \end{bmatrix} \tag{3.97}$$

where $Q_{11} = \Gamma_1^* F_1$. Let G be the matrix

$$G = \begin{bmatrix} I & M'V^{-1}M \\ 0 & I \end{bmatrix} \begin{bmatrix} M'V^{-1}M & 0 \\ 0 & \Phi^{-1}H(\Phi')^{-1} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ M'V^{-1}M & I \end{bmatrix} \quad (3.98)$$

which is Hermitian, nonnegative definite. Then by (3.50, 3.95),

$$J_1^* Q_{11} J_1 - Q_{11} = [F_1^* \quad \Gamma_1^*] G \begin{bmatrix} F_1 \\ \Gamma_1 \end{bmatrix} \geq 0, \quad (3.99)$$

and therefore[1] $Q_{11} \geq 0$ since the eigenvalues of J_1 are greater than one in absolute value. Thus Q and P_e are also nonnegative definite.

→ Let P_e be Hermitian, nonnegative definite. By (3.10),

$$(DP_e - C)^* P_e (DP_e - C) - P_e = R \quad (3.100)$$

where R denotes the matrix

$$R = [P_e^* \quad I] \begin{bmatrix} DB & D^*A \\ -C^*B-I & -C^*A \end{bmatrix} \begin{bmatrix} P_e \\ I \end{bmatrix} \quad (3.101)$$

By (3.50, 3.98), R is also equal to

$$R = [P_e^* \quad I] \; G \begin{bmatrix} P_e \\ I \end{bmatrix} \quad (3.102)$$

which implies that it is nonnegative definite. By (3.20), the matrix DP_e-C is related to J by the equation $(DP_e - C)\Gamma = \Gamma J$. Therefore let $\underline{\gamma}$ be an eigenvector of DP_e-C and let ψ be the corresponding eigenvalue. Then by pre- and postmultiplying

[1]See Appendix B.

(3.100) by $\underline{\gamma}^*$ and $\underline{\gamma}$, the equation

$$(\psi^*\psi-1)\underline{\gamma}^* P_{e-} = \underline{\gamma}^* R\underline{\gamma} \tag{3.103}$$

is obtained. This equation implies the conclusion in one of three ways:

(a) If $\underline{\gamma}^* P_e \underline{\gamma} > 0$ and $\underline{\gamma}^* R\underline{\gamma} > 0$, then

$$\psi^*\psi = 1 + \frac{\underline{\gamma}^* R\underline{\gamma}}{\underline{\gamma}^* P_e \gamma} \tag{3.104}$$

which implies that $|\psi| > 1$.

(b) If $\underline{\gamma}^* P_e \underline{\gamma} > 0$ and $\underline{\gamma}^* R\underline{\gamma} = 0$, then $|\psi| = 1$ and the equations

$$M' V^{-1} M P_e \underline{\gamma} = \underline{0} \tag{3.105}$$

$$H(\Phi')^{-1}\underline{\gamma} = \underline{0} \tag{3.106}$$

are implied by (3.98, 3.102). Also since $\underline{\gamma}$ is an eigenvector of DP_e-C, (3.50, 3.105, 3.106) imply the equations

$$\Phi'\underline{\gamma} = (\psi^*)^{-1}\underline{\gamma} \tag{3.107}$$

$$H\underline{\gamma} = \underline{0}, \tag{3.108}$$

so the random process has a random walk mode which is undriven.

(c) If $\underline{\gamma}^* P_e \underline{\gamma} = \underline{\gamma}^* R\underline{\gamma} = 0$, then $|\psi|$ may have any value, but (3.105-3.108) still apply. Thus $(\psi^*)^{-1}$ is equal to the eigenvalue of some mode of the random process which is undriven. Q.E.D.

If R_p is derogatory, then F_2 in (3.95) is not necessarily zero, so the conditions of the previous theorem become necessary but not sufficient for P_e to be nonnegative definite. However, if all of the eigenvalues in J are greater than one in absolute value, then P_e is nonnegative definite because $Q = Q_{11}$ whose definiteness does not depend on R_p being nonderogatory.

Some of the exceptions to the previous theorems that occur when R_p is derogatory are illustrated by the following numerical example.

Example 3.1:

The random process has a stable mode ($\lambda_1 = 0.5$) which is driven and unobservable and an unstable mode ($\lambda_2 = 2$) which is undriven and observable. The Φ, H, M, and V matrices are

$$\Phi = \begin{bmatrix} 2 & -1 \\ 0 & \frac{1}{2} \end{bmatrix} \qquad H = \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix} \tag{3.109a,b}$$

$$M = [-1 \quad 2/3] \qquad V = 1, \tag{3.109c,d}$$

and the R_p matrix is

$$R_p = \begin{bmatrix} 2 & -1 & 8 & 12 \\ 0 & 1/2 & 12 & 18 \\ 1/2 & -1/3 & 1/2 & 0 \\ -1/3 & 2/9 & 1 & 2 \end{bmatrix} . \tag{3.110}$$

The Jordan form of R_p is

$$\Omega = \text{diag}(2,\ 2,\ 1/2,\ 1/2), \tag{3.111}$$

and the corresponding matrix of eigenvectors is

$$E = \begin{bmatrix} 8 & 25 & 0 & 2 \\ 12 & 24 & 0 & 3 \\ 0 & 3 & 3 & 0 \\ 1 & 0 & -2 & 0 \end{bmatrix}, \tag{3.112}$$

and since there are two blocks in Ω associated with each distinct eigenvalue, R_p is derogatory. The permissible J matrices are $J_a = \text{diag}(2,\ 2)$, $J_b = \text{diag}(2,\ 1/2)$, and $J_c = \text{diag}(1/2,\ 1/2)$.

Even though R_p is derogatory, all of the solutions of $R_pT = TJ_a$ are column equivalent, so there is only one equilibrium solution,

$$P_e = \begin{bmatrix} 25/3 & 8 \\ 8 & 12 \end{bmatrix}, \tag{3.113}$$

corresponding to J_a. As expected, this solution is real and symmetric since J_a meets the conditions of Theorems 3.15 and 3.16, and it is nonnegative definite since both eigenvalues in J_a are greater than one in absolute value.

The solutions of $R_pT = TJ_b$ are not all column equivalent, so therefore there is family of equilibrium solutions,

$$P_e = \begin{bmatrix} 16/3 & 8 \\ 8 & 12 \end{bmatrix} + x \begin{bmatrix} 2/3 & 0 \\ 1 & 0 \end{bmatrix} + y \begin{bmatrix} 2/3 & 1 \\ 0 & 0 \end{bmatrix} - \frac{1}{9}xy \begin{bmatrix} 4/3 & 2 \\ 2 & 3 \end{bmatrix}, \tag{3.114}$$

corresponding to J_b. These solutions exist even though there is an eigenvalue in J_b equal to the eigenvalue of an unobservable mode, and some of these solutions, specifically

$$P_e = \begin{bmatrix} \frac{16}{3} + \frac{4}{3}x - \frac{4}{27}x^2 & 8 + x - \frac{2}{9}x^2 \\ 8 + x - \frac{2}{9}x^2 & 12 - \frac{1}{3}x^2 \end{bmatrix} \tag{3.115}$$

are real symmetric even though J_b does not satisfy the conditions of Theorems 3.15 and 3.16. The Q matrix corresponding to solutions (3.115) is

$$Q = \begin{bmatrix} 27(36 + x^2) & 27x \\ 27x & 0 \end{bmatrix} \tag{3.116}$$

which shows that although Q_{11} is positive, the Q matrix as a whole is indefinite because the $\Gamma_1^* F_2$ and $\Gamma_2^* F_1$ terms are nonzero. Thus although J_b satisfies the conditions of Theorem 3.17, the corresponding P_e matrices are not necessarily nonnegative definite.

The solutions of $R_p T = TJ_c$ are all column equivalent to a matrix whose Γ part is singular, so there is no corresponding equilibrium solution. This case, therefore, conforms to

Theorem 3.13.

The results obtained thus far indicate the following general conclusions: If there are no random walk modes in the random process which are either undriven or unobservable, then half of the eigenvalues of R_p are inside the unit circle and half are outside. Thus one of the permitted J matrices has eigenvalues which are all greater than one in absolute value, and the corresponding equilibrium matrix, if it exists, is symmetric nonnegative definite. The existence of this matrix is quite important since, as will be shown in Section 4, this is the only stable equilibrium. Theorem 3.13 implies that this matrix does not exist if there are any unstable modes in the random process which are unobservable, but does not imply that it exists if there are no such modes. However by taking advantage of the fact that the eigenvalues of J are outside of the unit circle, it can be shown that this condition is both necessary and sufficient. Theorem 3.14 implies that this equilibrium matrix is positive semidefinite if the random process has a stable mode which is undriven, and examples indicate that it is positive definite if there are no such modes. Theorem 3.17 implies that other nonnegative definite equilibrium solutions are possible only if the random process has one or more unstable modes which are undriven. These solutions are obtained by replacing the eigenvalue of the unstable mode by its conjugate reciprocal in J. Theorem 3.14 implies that these solutions are positive semidefinite. It is therefore expected that the solution corresponding to the J matrix whose eigenvalues are all outside the unit circle should be "more positive" than any of these solutions, and this is shown to be true in Theorem 3.19.

If the random process has a simple random walk mode which is undriven but observable, then R_p has a 2-by-2 Jordan block whose eigenvalue lies on the unit circle. The J matrix can therefore have at most n-1 eigenvalues outside the unit circle and one on the unit circle. The corresponding equilibrium solution is real and symmetric since J satisfies Theorems 3.15 and 3.16, and by Theorems 3.14 and 3.17 it is positive semidefinite. On the other hand, if the random process has a simple random walk mode which is driven and unobservable, then R_p again has a 2-by-2 Jordan block whose eigenvalue lies on the unit circle, but by Theorem 3.13 there is no equilibrium solution corresponding to the J matrix which has one eigenvalue on, and the rest outside of, the unit circle. When there is a simple random walk mode which is both undriven and unobservable, then R_p is derogatory with two 1-by-1 Jordan blocks associated with an eigenvalue on the unit circle, and there is a continuous spectrum of equilibrium matrices which can be positive definite, positive semidefinite, or indefinite. When the random process has a multiple random walk mode which is undriven and/or unobservable, the situation becomes quite complex and is best studied by direct solution of the quadratic matrix equation.

The following theorem establishes necessary and sufficient conditions for Γ to be singular when the eigenvalues in J are outside of the unit circle.

Theorem 3.18:

If the system has no undriven or unobservable random walk modes such that the equation

$$\begin{bmatrix} H(\Phi')^{-1}M'V^{-1}M + \Phi & H(\Phi')^{-1} \\ (\Phi')^{-1}M'V^{-1}M & (\Phi')^{-1} \end{bmatrix} \begin{bmatrix} F \\ \Gamma \end{bmatrix} = \begin{bmatrix} F \\ \Gamma \end{bmatrix} J \qquad (3.117)$$

has a solution with linearly independent columns when all the eigenvalues in J are greater than one in absolute value, then Γ is singular if and only if the system has an unstable mode which is unobservable.

Proof:

$\Leftarrow$ If the system has an unstable mode which is unobservable, then there exists a vector $\underline{v}$ such that $\Phi\underline{v} = \lambda\underline{v}$ with $|\lambda| > 1$, and $M\underline{v} = \underline{0}$. Therefore $\underline{v}^*(\Phi')^{-1} = (\lambda^*)^{-1}\underline{v}^*$ and by the lower half of (3.117)

$$\underline{v}^*\Gamma(J-(\lambda^*)^{-1}I) = \underline{0} \; . \qquad (3.118)$$

Since the eigenvalues of J are outside of and $(\lambda^*)^{-1}$ is within the unit circle, $J - (\lambda^*)^{-1}I$ is nonsingular, so $\underline{v}^*$ is a null row vector of Γ.

→ Let $\underline{n}$ be a null vector of Γ. From the lower half of (3.117)

$$(M'V^{-1}MF + \Gamma)J^{-1} = \Phi'\Gamma \ , \tag{3.119}$$

and by premultiplying by $\underline{n}^*(J^*)^{-1}F^*$ and postmultiplying by $\underline{n}$, the equation

$$\underline{n}^*(J^*)^{-1}F^*(M'V^{-1}M)FJ^{-1}\underline{n} + \underline{n}^*(J^*)^{-1}(F^*\Gamma)J^{-1}\underline{n} = 0 \tag{3.120}$$

is obtained. Theorems 3.15, 3.16, and 3.17 imply that $F^*\Gamma$ is real, symmetric, and nonnegative definite since all of the eigenvalues in J are outside of the unit circle. Thus since $M'V^{-1}M$ is also nonnegative definite, both terms in (3.120) must be zero, so $MFJ^{-1}\underline{n} = \underline{0}$. Thus by postmultiplying (3.119) by $\underline{n}$, it can be seen that $J^{-1}\underline{n}$ is also a null vector of Γ. This argument can be repeated indefinitely to show that every vector of the form $J^{-i}\underline{n}$ is a null vector of both Γ and MFJ^{-1}. Thus by Appendix B, there is a vector $\underline{v}\varepsilon\{\underline{n}, J^{-1}\underline{n}, \ldots\}$ such that $J^{-1}\underline{v} = \lambda\underline{v}$ with $|\lambda| < 1$, $\Gamma\underline{v} = \underline{0}$, and $MF\underline{v} = \underline{0}$. Now by postmultiplying the top part of (3.117) by $J^{-1}\underline{v}$, the equation

$$\Phi(F\underline{v}) = \lambda^{-1}(F\underline{v}) \tag{3.121}$$

is obtained. Finally, $F\underline{v} \neq 0$ since the columns of $\begin{bmatrix} F \\ \Gamma \end{bmatrix}$ are linearly independent. Thus there is an unstable mode which is unobservable because $F\underline{v}$ is and eigenvector of Φ, whose corresponding eigenvalue is outside of the unit circle, and it is a null vector of M. Q.E.D.

The following theorem shows that the effect of replacing an eigenvalue of J, which is equal to the eigenvalue of an undriven unstable mode, by its conjugate reciprocal is to make the corresponding equilibrium solution less positive definite.

<u>Theorem 3.19</u>:

If ψ_1 is equal to the eigenvalue of an unstable mode which is undriven and observable, and if P_{ea} and P_{eb} are Hermitian nonnegative definite equilibrium solutions of the covariance equation associated with the Jordan matrices

$$J_a = \text{diag}(\psi_1, \psi_2, \ldots, \psi_n) \tag{3.122}$$

$$J_b = \text{diag}((\psi_1^*)^{-1}, \psi_2, \ldots, \psi_n) \tag{3.123}$$

respectively, then $P_{ea} \geq P_{eb}$.

<u>Proof</u>:

The equilibrium solutions P_{ea} and P_{eb} are equal to

$$P_{ea} = F_a \Gamma_a^{-1} = [\underline{f}_a, \underline{f}_2, \ldots, \underline{f}_n][\underline{\gamma}_a, \underline{\gamma}_2, \ldots, \underline{\gamma}_n]^{-1} \tag{3.124}$$

$$P_{eb} = F_b \Gamma_b^{-1} = [\underline{f}_b, \underline{f}_2, \ldots, \underline{f}_n][\underline{\gamma}_b, \underline{\gamma}_2, \ldots, \underline{\gamma}_n]^{-1} \tag{3.125}$$

respectively where the vectors $\underline{f}_i$ and $\underline{\gamma}_i$ satisfy the equation

$$R_p \begin{bmatrix} \underline{f}_a & \underline{f}_b & \underline{f}_2 & \cdots & \underline{f}_n \\ \underline{\gamma}_a & \underline{\gamma}_b & \underline{\gamma}_2 & \cdots & \underline{\gamma}_n \end{bmatrix}$$

$$= \begin{bmatrix} \underline{f}_a & \underline{f}_b & \underline{f}_2 & \cdots & \underline{f}_n \\ \underline{\gamma}_a & \underline{\gamma}_b & \underline{\gamma}_2 & \cdots & \underline{\gamma}_n \end{bmatrix} \mathrm{diag}(\psi_1, (\psi_1^*)^{-1}, \psi_2, \ \ldots \ \psi_n) . \qquad (3.126)$$

Therefore

$$\Gamma_a^*(P_{ea}-P_{eb})\Gamma_a = \Gamma_a^*\{[\underline{f}_a, \underline{f}_2, \ldots, \underline{f}_n]$$

$$-[\underline{f}_b, \underline{f}_2, \ldots, \underline{f}_n][\gamma_b, \gamma_2, \ldots, \gamma_n]^{-1}[\gamma_a, \gamma_2, \ldots, \gamma_n]\}$$

$$= \begin{bmatrix} \underline{\gamma}_a^* \\ \underline{\gamma}_2^* \\ \cdot \\ \cdot \\ \underline{\gamma}_n^* \end{bmatrix} [\underline{f}_a - P_{eb}\underline{\gamma}_a, \underline{0}, \ldots, \underline{0}] = \begin{bmatrix} d & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \ldots & 0 \end{bmatrix} \qquad (3.127)$$

where $d = \underline{\gamma}_a^*(\underline{f}_a - P_{eb}\underline{\gamma}_a)$. By (1.25, 1.28e),

$$P_{eb} = \Phi P_{eb}(I + M'V^{-1}MP_{eb})^{-1}\Phi' + H , \qquad (3.128)$$

and by (3.126)

$$\psi_1\underline{\gamma}_a = (\Phi')^{-1}(M'V^{-1}M\underline{f}_a + \underline{\gamma}_a) \qquad (3.129)$$

$$\psi_1\underline{f}_a = \psi_1 H\underline{\gamma}_a + \Phi\underline{f}_a . \qquad (3.130)$$

Thus[1]

$$(\psi_1^* \psi_1 - 1)d$$

$$= (\underline{f}_a^* M' V^{-1} M + \underline{\gamma}_a^*)\{\underline{f}_a - P_{eb}(M\ V^{-1} M P_{eb} + I)^{-1}(M\ V^{-1} M \underline{f}_a + \underline{\gamma}_a)\}$$

$$- \underline{\gamma}_a^* \underline{f}_a + \underline{\gamma}_a^* P_{eb} \underline{\gamma}_a$$

$$= (\underline{f}_a^* - \underline{\gamma}_a^* P_{eb}) M' (M P_{eb} M' + V)^{-1} M (\underline{f}_a - P_{eb} \underline{\gamma}_1) \ . \quad (3.131)$$

The right side of (3.131) is nonnegative because $P_{eb} \geq 0$ and $V > 0$.[2] Therefore since $\psi_1^* \psi_1 - 1$ is positive, d is also nonnegative, which by (3.127) implies that $P_{ea} \geq P_{eb}$. Q.E.D.

The final theorem in this section demonstrates that any nonnegative definite equilibrium solution is singular when the random process has a stable mode which is undriven. This fact is implied by Theorems 3.14 and 3.17 when R_p is nonderogatory, but as the following theorem shows, it is not restricted to such cases.

Theorem 3.20:

If P_e is nonnegative definite and the system has a stable mode which is undriven, then P_e is singular.

---

[1] The matrix inversion lemma and the identity $M'(MPM' + V)^{-1} M = M'V^{-1}M - M'V^{-1}MP(M'V^{-1}MP + I)^{-1} M'V^{-1} M$ must be used to obtain the final form of (3.131).

[2] Numerical examples seem to indicate that the matrix $(MP_eM'+V)$ is positive definite for all equilibrium solutions of the covariance equation. If this can be shown to be true in general, then the restriction that ψ_1 be equal to the eigenvalue of an undriven unstable mode can be dropped.

<u>Proof</u>:

By Appendix C there exists a vector $\underline{n}_d$ such that $\Phi'\underline{n}_d = \lambda_d^*\underline{n}_d$ with $|\lambda_d| < 1$ and $H\underline{n}_d = \underline{0}$. Therefore

$$\underline{n}_d^* P_e \underline{n}_d = \underline{n}_d^* [\Phi[P_e - P_e M'(MP_eM' + V)^{-1}MP_e]\Phi + H\}\underline{n}_d$$

$$= (\lambda_d^* \lambda_d)\underline{n}_d^* [P_e - P_e M'(MP_eM' + V)^{-1}MP_e]\underline{n}_d \tag{3.132}$$

and

$$(1-\lambda_d^*\lambda_d)\underline{n}_d^* P_e \underline{n}_d = -(\lambda_d^*\lambda_d)\underline{n}_d^* P_e M'(MP_eM' + V)^{-1}MP_e\underline{n}_d \,. \tag{3.133}$$

Now since $|\lambda_d| < 1$ and $P_e \geq 0$, the left side of (3.142) is nonnegative and the right side is nonpositive, so both sides are zero. Thus P_e is singular and $\underline{n}_d$ is one of its null vectors.

d. <u>The a posteriori equilibrium matrices</u> All of the results derived thus far in Section 3 apply to the a priori equilibrium matrices, P_e. The a posteriori equilibrium matrices, Q_e, and their properties can be derived from the equations:

$$Q_e = P_e - P_e M'(MP_eM' + V)^{-1}MP_e \tag{3.134a}$$

$$= P_e(M'V^{-1}MP_e + I)^{-1} \tag{3.134b}$$

$$= (P_eM'V^{-1}M + I)^{-1}P_e \tag{3.134c}$$

$$= (I-K_eM)P_e(I-M'K_e') + K_eVK_e' \tag{3.134d}$$

and

$$P_e = \Phi Q_e \Phi' + H \tag{3.135}$$

where $K_e = P_e M'(MP_e M' + V)^{-1}$. Equations (3.134a, 3.135) imply that Q_e is symmetric if and only if P_e is symmetric, (3.134d, 3.135) imply that Q_e is nonnegative definite if and only if P_e is nonnegative definite, (3.134c) implies that the null spaces of Q_e and P_e are identical, and (3.134a) implies that $Q_e \leq P_e$ whenever $MP_e M' + V$ is positive definite.

A quadratic matrix equation for Q_e can be obtained from (3.134b) by postmultiplying by $(M'V^{-1}MP_e + I)$ and replacing P_e by the right side of (3.135):

$$H(\Phi')^{-1} + \Phi Q_e - Q_e(M'V^{-1}MH+I)(\Phi')^{-1} - Q_e M'V^{-1}M\Phi Q_e = 0. \tag{3.136}$$

The solutions of this equation are given by $Q_e = F_q \Gamma_q^{-1}$ where

$$R_q \begin{bmatrix} F_q \\ \Gamma_q \end{bmatrix} = \begin{bmatrix} F_q \\ \Gamma_q \end{bmatrix} J_q \tag{3.137}$$

and

$$R_q = \begin{bmatrix} \Phi & H(\Phi')^{-1} \\ M'V^{-1}M\Phi & (M'V^{-1}MH + I)(\Phi')^{-1} \end{bmatrix} \tag{3.138}$$

which can be factored into the product

$$R_q = \begin{bmatrix} I & 0 \\ M'V^{-1}M & I \end{bmatrix} \begin{bmatrix} I & H \\ 0 & I \end{bmatrix} \begin{bmatrix} \Phi & 0 \\ 0 & (\Phi')^{-1} \end{bmatrix} . \tag{3.139}$$

Also, R_p and R_q are related by the equation

$$R_q = \begin{bmatrix} I & 0 \\ M'V^{-1}M & I \end{bmatrix} R_p \begin{bmatrix} I & 0 \\ -M\ V^{-1}M & I \end{bmatrix}, \tag{3.140}$$

so they are similar and have the same Jordan matrix, Ω. Furthermore, if P_e and Q_e are related by (3.135), then the Jordan matrices J_p and J_q are similar. This results from the identity

$$(D_pP_e - C_p)\Gamma_p = \Gamma_pJ_p , \tag{3.141}$$

which by (3.50, 3.135) implies the equation

$$[M'V^{-1}M(\Phi Q_e\Phi' + H) + I]\Gamma_p = \Phi'\Gamma_pJ_p , \tag{3.142}$$

and from the identity

$$(D_qQ_e - C_q)\Gamma_q = \Gamma_qJ_q , \tag{3.143}$$

which by (3.138) implies the equation

$$[M'V^{-1}M(\Phi Q_e\Phi' + H) + I](\Phi')^{-1}\Gamma_q = \Gamma_qJ_q . \tag{3.144}$$

Therefore by (3.142, 3.144),

$$(\Gamma_q^{-1}\Phi'\Gamma_p)J_p = J_q(\Gamma_q^{-1}\Phi'\Gamma_p) , \tag{3.145}$$

which implies that J_p and J_q are similar. In fact, J_p and J_q are identical if their eigenvalues are arranged in the same order.

4. Stability of the time-invariant covariance equation

a. Local stability In this section, the behavior of the covariance matrix in a suitably small region around an equilibrium solution will be examined. The covariance equation can be converted to a more suitable form for this purpose by following the same procedure as was used to obtain (3.49). The resulting equation is

$$A + BP_k + P_{k+1}C - P_{k+1}DP_k = 0 \,, \tag{3.146}$$

where A, B, C, and D are the n-by-n matrices defined by (3.50). Since this involves a postmultiplication by $I + M'V^{-1}MP_k$, there may be sequences which satisfy the above equation which do not satisfy the original covariance equation due to $I + M'V^{-1}MP_k$ being singular for some value of k. However, the following theorem shows that this can happen only if the sequence which satisfies the original covariance equation terminates due to a singular $MP_kM' + V$ matrix.

Theorem 3.21:

If V is nonsingular, then $MPM' + V$ is singular if and only if $I + M'V^{-1}MP$ is singular.

Proof:

The proof is based on the identity

$$(I + M'V^{-1}MP)M' = M'V^{-1}(MPM' + V) \quad . \tag{3.147}$$

$\rightarrow$Let $\underline{x}$ be a null vector of $MPM' + V$ and let $\underline{y} = M'\underline{x}$. Then $MP\underline{y} + V\underline{x} = \underline{0}$, and since V is nonsingular, $\underline{y}$ cannot be $\underline{0}$. The conclusion is obtained by postmultiplying (3.147) by $\underline{x}$:

$$(I + M'V^{-1}MP)\underline{y} = M'V^{-1}(MPM' + V)\underline{x} = \underline{0}. \tag{3.148}$$

$\leftarrow$Let $\underline{v}'$ be a null row vector of $I + M'V^{-1}MP$ and let $\underline{w}' = \underline{v}'M'V^{-1}$. Then $\underline{v}' + \underline{w}'MP = \underline{0}'$, which implies that $\underline{w}$ cannot be $\underline{0}$. The conclusion is obtained by premultiplying (3.147) by $\underline{v}'$:

$$\underline{w}'(MPM' + V) = \underline{v}'(I + M'V^{-1}MP)M' = \underline{0}'. \tag{3.149}$$

Q.E.D.

Also since $\det(I + M'V^{-1}MP_k)$ is a continuous function of the elements of P_k and since $I + M'V^{-1}MP_e$ is nonsingular for every equilibrium solution of the covariance equation, there must be a region around every P_e for which the matrix $I + M'V^{-1}MP_k$ is nonsingular. Thus (3.146) is suitable for studying the local stability of an equilibrium solution.

Now let P_k be replaced by $P_e + U_k$ in (3.146). Then U_k satisfies the difference equation

$$U_{k+1} = (B - P_eD)U_k[(DP_e-C) + DU_k]^{-1} . \qquad (3.150)$$

Intuitively it can be seen that if U_k is small enough, then the stability of (3.150) is determined by the stability of the linearized equation

$$U_{k+1} = (B-P_eD)U_k(DP_e-C)^{-1} . \qquad (3.151)$$

Since the Jordan forms of $B-P_eD$ and DP_e-C are K and J respectively, and since the eigenvalues in K are the conjugate reciprocals of the eigenvalues in J when P_e is Hermitian, (3.151) is stable if all of the eigenvalues in J are outside of the unit circle, and it is unstable if any eigenvalue in J is inside the unit circle. The following two theorems show that (3.150) is stable or unstable under the same conditions.

<u>Theorem 3.22</u>:

If all of the eigenvalues in the J matrix associated with an equilibrium solution are outside of the unit circle, then any P_k matrix starting sufficiently near P_e approaches P_e as $k\to\infty$.

<u>Proof</u>:

Let $U_k = L_kR_k$. Then L_k and R_k satisfy the difference equations

$$L_{k+1} = (B-P_eD)L_k \qquad (3.152)$$

$$R_{k+1} = R_k[(DP_e - C) + DU_k]^{-1} \;. \tag{3.153}$$

Let S be a symmetric, positive definite matrix such that $T = (DP_e - C)S(DP_e - C)^* - S > 0$, let the norm of a vector be defined as $||x||^2 = \underline{x}^* S \underline{x}$, and let the norm of R_k be defined as

$$||R_k|| \triangleq \max_{||\underline{x}||=1} ||R_k^* \underline{x}|| \;. \tag{3.154}$$

Let $\underline{y} = [(DP_e - C)^* + U_k^* D^*]^{-1} R_k^* \underline{x}$; then

$$[(DP_e - C)^* + U_k^* D^*]\underline{y} = R_k^* \underline{x} \;. \tag{3.155}$$

Note that

$$\begin{aligned} ||(DP_e - C)^* \underline{y}||^2 &= \underline{y}^* (DP_e - C) S (DP_e - C)^* \underline{y} \\ &= \underline{y}^* S \underline{y} + \underline{y}^* T \underline{y} \;, \end{aligned} \tag{3.156}$$

so therefore

$$||(DP_e - C)^* y|| \geq \alpha ||y|| \tag{3.157}$$

where $\alpha > 1$. Also

$$U_k^* D^* \underline{y} = R_k^* L_0^* (B^* - D^* P_e^*)^k D^* \underline{y} \tag{3.158}$$

and

$(B^* - D^* P_e^*)^k \to 0$ as $k \to \infty$, so

$$||U_k^* D^* \underline{y}|| \leq \gamma \cdot ||R_k|| \cdot ||\underline{y}|| \tag{3.159}$$

where $\gamma = \beta ||L_0|| \cdot ||D||$ and β is a positive number such that

$||(B-P_eD)^k|| \leq \beta$ for all k. Therefore if

$$||R_k|| < \frac{\alpha-1}{\gamma}, \tag{3.160}$$

then

$$||[(DP_e-C)^* + U_k^* D^*]\underline{y}|| \geq (\alpha-\gamma||R_k||)||\underline{y}|| = \delta||\underline{y}|| \tag{3.161}$$

where $\delta > 1$, and by (3.155)

$$||y|| < \delta^{-1} ||R_k^* \underline{x}|| \tag{3.162}$$

for all $\underline{x}$. But

$$||R_{k+1}|| = \max_{||\underline{x}||=1} ||\underline{y}||, \tag{3.163}$$

so

$$||R_{k+1}|| \leq \delta^{-1} ||R_k|| . \tag{3.164}$$

Thus both L_k and R_k approach 0 as $k\to\infty$, so U_k also vanishes. This proves the convergence of P_k to P_e. Note that if $||U_0|| \triangleq ||L_0||\cdot||R_0||$, then by (3.160)

$$||U_0|| < \frac{\alpha-1}{\beta||D||} .$$

Theorem 3.23:

If the J matrix associated with an equilibrium solution has an eigenvalue within the unit circle, then there are P_k matrices starting arbitrarily near P_e which diverge from P_e.

<u>Proof</u>:

By utilizing the identity[1]

$$(DP_e - C)^{-1} = (B - P_eD)^* , \tag{3.165}$$

Equation (3.150) can be written as

$$U_{k+1} = (B-P_eD)U_k(B-P_eD)^*[I+DU_k(B-P_eD)^*]^{-1} . \tag{3.166}$$

Under the conditions of the hypothesis, $B-P_eD$ has an eigenvalue, μ, outside of the unit circle. Let $\underline{v}$ be the corresponding eigenvector with $\underline{v}^*\underline{v} = 1$. Let $U_0 = \delta^2\underline{vv}^*$ where δ is a real, arbitrarily small, positive number. Then $U_k = \underline{\ell}_k\underline{r}_k^*$ where

$$\underline{\ell}_k = \delta\mu^k\underline{v} \tag{3.167}$$

and

$$\underline{r}_{k+1} = [I + (B-P_eD)U_k^*D^*]^{-1}(B-P_eD)\underline{r}_k \tag{3.168}$$

with $\underline{r}_0 = \delta\underline{v}$. Suppose

$$\underline{r}_k = \frac{\delta\mu^k}{1 + \delta^2\mu\alpha(1 + |\mu|^2 + \ldots + |\mu|^{2k-2})}\,\underline{v} , \tag{3.169}$$

where $\alpha = \underline{v}^*D^*\underline{v}$. Then insertion of (3.167, 3.169) into (3.168) results in the equation

---

[1]This identity is derived as follows:

$$(B-P_eD)^* = \Phi - M'V^{-1}M\Phi^{-1}(P_e-H) = [I-M'V^{-1}MP_e(M'V^{-1}MP_e+I)^{-1}]\Phi$$

$$= (M'V^{-1}MP_e + I)^{-1}\Phi' = (DP_e-C)^{-1}\Phi.$$

$$\underline{r}_{k+1} = \frac{\delta\mu^{k-1}}{1 + \delta^2\mu\alpha(1 + |\mu|^2 + \ldots + |\mu|^{2k-2})}\,\underline{w} \quad (3.170)$$

where

$$\underline{w} = [I + \frac{\delta^2\mu^{k+1}(\mu^*)^k}{1 + \delta^2\mu\alpha(1 + |\mu|^2 + \ldots + |\mu|^{2k-2})}\,\underline{v}\,\underline{v}^* D^*]^{-1}\underline{v}\ . \quad (3.171)$$

Premultiplication by the bracketed term in (3.71) results in the equation

$$\underline{w} + \frac{\delta^2\mu|\mu|^{2k}(\underline{v}^* D^* \underline{x})}{1 + \delta^2\mu\alpha(1 + |\mu|^2 + \ldots + |\mu|^{2k-2})}\,\underline{v} = \underline{v} \quad (3.172)$$

which implies that

$$\underline{w} = \frac{1 + \delta^2\mu\alpha(1 + |\mu|^2 + \ldots + |\mu|^{2k-2}}{1 + \delta^2\mu\alpha(1 + |\mu|^2 + \ldots + |\mu|^{2k})}\,\underline{v} \quad (3.173)$$

and

$$r_{k+1} = \frac{\delta\mu^{k+1}}{1 + \delta^2\mu\alpha(1 + |\mu|^2 + \ldots + |\mu|^{2k})}\,\underline{v}\ . \quad (3.174)$$

Furthermore, a similar analysis shows that (3.169) is true for k = 1, so by induction it is true for all $k \geq 1$. Thus $U_k = \beta_k \underline{v}\,\underline{v}^*$ where

$$\beta_k = \frac{\delta^2|\mu|^{2k}}{1 + \dfrac{\delta^2\mu^*\alpha^*(|\mu|^{2k}-1)}{|\mu|^2 - 1}}\ . \quad (3.175)$$

Instability can now be proved by observing that: (a) if $\alpha \neq 0$ and

$$n \geq \frac{\ln\left(\frac{|\mu|^2-1}{10\delta^2|\mu\alpha|}\right)}{2\ln|\mu|}, \tag{3.176}$$

then

$$|\beta_k| \geq \frac{|\mu|^2 - 1}{11\cdot|\mu\alpha|}, \tag{3.177}$$

or (b) if $\alpha = 0$ and $k \geq \frac{\ln(\varepsilon/\delta^2)}{2\ln|\mu|}$ then $|\beta_k| \geq \varepsilon$. Q.E.D.

If J has one or more eigenvalues on the unit circle, with the remainder outside of the unit circle, then there is no particular advantage in making a local stability analysis because the nearness of P_k to P_e does not make the analysis any simpler than a global stability analysis.

b. <u>Global stability</u> In this section, a discrete analogue of Potter's [13] global stability theorem is stated and proved. The conditions imposed on the random process in this theorem are probably the most realistic in terms of an actual implementation of the Kalman filter, but in terms of a stability analysis, a relaxation of these conditions leads to more interesting properties. The author has investigated several ways of analyzing the stability of the covariance equation under such conditions. The results of this investigation and some speculations based on numerical examples are stated after the proof of the following theorem.

Theorem 3.24:

If all unstable and random walk modes are both driven and observable and if P_0 is symmetric nonnegative definite, then P_k approaches the P_e whose corresponding J matrix has all of its eigenvalues outside of the unit circle.

Proof:

(a) Suppose $P_0 \geq P_e$ and $P_k = P_e + U_k$. Then by Theorem 3.4, $U_k \geq 0$ and it satisfies the equation

$$U_{k+1} = (B-P_eD)U_k[(DP_e-C) + DU_k]^{-1} . \tag{3.178}$$

Let $W_0 = U_0$ and

$$W_{k+1} = (B-P_eD)W_k(B-P_eD)^* . \tag{3.179}$$

Since the Jordan form of $B-P_eD$ is K, whose eigenvalues are within the unit circle, $W_k \to 0$ as $k \to \infty$. Let $X_k = W_k - U_k$, then X_k satisfies the equation

$$X_{k+1} = (B-P_eD)\{X_k+U_k-U_k[I+(B-P_eD)^*DU_k]^{-1}\}(B-P_eD)^* , \tag{3.180}$$

and since

$$\begin{aligned} U_k - U_k[I+(B-P_eD)^*DU_k]^{-1} &= U_k-U_k(I+M'V^{-1}MP_k)^{-1}(I+M'V^{-1}MP_e) \\ &= U_k(I+M'V^{-1}MP_k)^{-1}M'V^{-1}MU_k \end{aligned}$$

$$= U_k(I+M'V^{-1}MP_k)^{-1}(M'V^{-1}M+M'V^{-1}MP_kM'V^{-1}M)$$

$$\cdot(I+P_kM'V^{-1}M)^{-1}U_k \tag{3.181}$$

which is nonnegative definite, $X_k \geq 0$ for all $k \geq 0$. Thus U_k is constrained by the inequalities

$$0 \leq U_k = W_k - X_k \leq W_k , \tag{3.182}$$

so it must also approach zero as $k\to\infty$.

(b) Suppose $P_0 = 0$, P_e is positive definite, and $P_k = P_e - U_k$. Then U_k satisfies the difference equation

$$U_{k+1} = (B-P_eD)U_k[(DP_e-C)-DU_k]^{-1} \tag{3.183}$$

with $U_0 = P_e$. Let[1] $W = (I + M'V^{-1}MP_e)^{-1}M'V^{-1}M$, let $X_0 = P_e^{-1}$, and let

$$X_{k+1} = X_k-[(B-P_eD)^k]^*W(B-P_eD)^k . \tag{3.184}$$

Then

$$X_k = P_e^{-1} - S_k \geq P_e^{-1} - S \tag{3.185}$$

where

$$S_k-(B-P_eD)^*S_k(B-P_eD) = W-[(B-P_eD)^k]^*W(B-P_eD)^k \tag{3.186}$$

---

[1]W is symmetric nonnegative definite since $(I + M'V^{-1}MP_e)^{-1}M'V^{-1}M = M'(MP_eM'+ V)^{-1}M$.

and

$$S-(B-P_eD)^*S(B-P_eD) = W . \tag{3.187}$$

Equation (3.187) implies that

$$(DP_e-C)(P_e^{-1}-S)(DP_e-C)^* - (P_e^{-1}-S)$$

$$= (\Phi')^{-1}(P_e^{-1}+ M'V^{-1}M)\Phi^{-1} - P_e^{-1}$$

$$= (P_e-H)^{-1} - P_e^{-1} , \tag{3.188}$$

which is nonnegative definite since the inequalities

$$P_e \geq P_e - H = \Phi(P_e^{-1} + M'V^{-1}M)^{-1}\Phi' > 0 \tag{3.189}$$

imply (Bellman [23], p. 92) that $(P_e-H)^{-1} \geq P_e^{-1}$. Therefore by (3.188) and Appendix B, $P_e^{-1}-S$ is at least nonnegative definite. Now suppose $\underline{v}$ is an eigenvector of $(DP_e-C)^*$ and a null vector of $(P_e-H)^{-1} - P_e^{-1}$. Then $P_e^{-1}\underline{v}$ is an eigenvector of Φ' and a null vector of H, which implies that there is an undriven mode. But this is impossible since by hypothesis, any undriven mode must be a stable mode, and by Theorem 3.20 when a stable mode is undriven, P_e is singular. Thus no eigenvector of $(DP_e-C)^*$ is a null vector of $(P_e-H)^{-1} - P_e^{-1}$, so $P_e^{-1} - S$ is positive definite, which by (3.185) implies that X_k is nonsingular for all k.

The expression

$$U_k = (B-P_eD)^kX_k^{-1}[(B-P_eD)^k]^* , \tag{3.190}$$

therefore, exists for all k, and it is set equal to U_k because: (1) it is equal to P_e when $k = 0$, and (2) it satisfies the difference equation for U_k since

$$\begin{aligned}
U_{k+1} &= (B-P_eD)^{k+1}\{X_k-[(B-P_eD)^k]^*W(B-P_eD)^k\}^{-1}[(B-P_eD)^{k+1}]^* \\
&= (B-P_eD)^{k+1}X_k^{-1}\{(DP_e-C) \\
&\quad - (DP_e-C)^{-k}D(B-P_eD)^kX_k^{-1}\}^{-1}(DP_e-C)^{-k} \\
&= (B-P_eD)^{k+1}X_k^{-1}[(B-P_eD)^k]^* \\
&\quad \cdot\{(DP_e-C) - D(B-P_eD)^kX_k^{-1}[(B-P_eD)^k]^*\}^{-1} \\
&= (B-P_eD)U_k[(DP_e-C) - DU_k]^{-1} . \qquad (3.191)
\end{aligned}$$

Finally (3.190) implies that $U_k \to 0$ as $k\to\infty$ since the eigenvalues of $B-P_eD$ are within the unit circle and X_k^{-1} is bounded by $(P_e^{-1}-S)^{-1}$.

(c) Now suppose P_e is positive semidefinite and $P_0=0$. Let the columns of L and N be basis vectors for $\mathcal{R}(P_e)$ and $\eta(P_e)$ respectively. Then since both terms in

$$N^*\Phi(P_eM'V^{-1}M + I)^{-1}P_e\Phi'N + N^*HN = N^*P_eN = 0 \quad (3.192)$$

are nonnegative definite, $P_e\Phi'N = 0$ and $HN = 0$. Thus $\eta(P_e)$ is invariant under Φ' and is contained in $\eta(H)$, and by the properties of orthogonal subspaces, $\mathcal{R}(P_e)$ is invariant under Φ and contains $\mathcal{R}(H)$. Thus if L is chosen such that $P_e = LL^*$, then $\Phi L = L\Phi_\ell$ and $H = LGL^*$ where $G \geq 0$. Therefore if $X_0=0$ and

$$X_{k+1} = \Phi_\ell[X_k - X_k L^* M'(MLX_k L^* M' + V)^{-1} MLX_k]\Phi_\ell^* + G, \quad (3.193)$$

then $P_k = LX_k L^*$.

Note that $\dim(X_k) = \operatorname{rank}(P_e) < \dim(P_k)$, and that (3.193) is a covariance equation whose equilibrium solution is $X_e = I$. Furthermore, no unstable or random walk mode in (3.193) is unobservable since if $\underline{v}$ were a vector such that $\Phi_\ell \underline{v} = \lambda \underline{v}$ with $|\lambda| \geq 1$ and $ML\underline{v} = \underline{0}$, then $\Phi(L\underline{v}) = \lambda(L\underline{v})$ and $M(L\underline{v}) = \underline{0}$ which implies that the original system has an unobservable random walk or unstable mode. Finally, all modes in (3.193) are completely driven. To show this let the matrix N_d be defined such that its columns are basis vectors for $\eta(R_d)$, $\Phi' N_d = N_d \Lambda_d$, and $HN_d = 0$, where $R_d = \sum_{i=0}^{n-1} \Phi^i H(\Phi^i)'$ and Λ_d is a Jordan matrix whose eigenvalues are within the unit circle since only stable modes may be undriven. Then the covariance equation implies that

$$N_d^* P_e N_d - \Lambda_d^* N_d^* P_e N_d \Lambda_d$$
$$+ \Lambda_d^* N_d^* P_e M'(MP_e M' + V)^{-1} MP_e N_d \Lambda_d = 0, \quad (3.194)$$

and since the magnitude of the eigenvalues in Λ_d is less than one, each column in N_d is a null vector of P_e. Thus $\eta(R_d) \subset \eta(P_e)$ and $\mathcal{R}(R_d) \supset \mathcal{R}(P_e)$. Now $R_d = LR_g L^*$ where

$R_g = \sum_{i=0}^{n-1} \Phi_\ell^{\,i} G(\Phi_\ell^{\,i})^*$ is the state covariance matrix for the reduced system, and therefore $\mathcal{R}(R_g) = \mathcal{R}(P_e)$ and R_g must be positive definite. Thus (3.193) satisfies the conditions imposed in the (b) part of this proof, so $X_k \to X_e$ and $P_k \to P_e$ as $k\to\infty$.

(d) Finally, let P_0 be any nonnegative definite matrix, let $P_a(t_0) = 0$, and let $P_b(t_0)$ be a matrix greater than or equal to both P_0 and P_e. Then by Theorem 3.4, P_k is constrained by the inequalities

$$P_a(t_k) \leq P_k \leq P_b(t_k) \quad , \tag{3.195}$$

and since both $P_a(t_k)$ and $P_b(t_k)$ approach P_e as $k\to\infty$, P_k also approaches P_e. Q.E.D.

When a random walk or unstable mode is observable but undriven, the results are much more interesting. Numerical studies indicate the following conclusions: (a) when a random walk mode is undriven and observable, then there is a single nonnegative definite P_e matrix which is singular, and if P_0 is any nonnegative definite matrix then $P_k \to P_e$, but there are nonpositive or indefinite values of P_0 which are arbitrarily close to P_e and yet cause P_k to diverge from P_e. (b) When an unstable mode is undriven and observable, then there are two nonnegative definite equilibrium solutions, P_{e1} which is positive definite and stable, and P_{e2} which is positive semidefinite and unstable. The only solutions which approach P_{e2} are those that start from an initial

P_0 matrix for which $P_0\underline{n} = \underline{0}$, where $\underline{n}$ is a null vector of P_{e2} and H and is an eigenvector of Φ' associated with the unstable mode. All other nonnegative definite values of P_0 cause $P_k \to P_{e1}$.

These conclusions are partially supported by the following analytic results. Suppose the random process has a simple random walk mode which is undriven and observable. Then Φ' has an eigenvector, $\underline{n}$, which is a null vector of P_e, and $(B-P_eD)$ has an eigenvector, $\underline{v}$, which is not orthogonal to $\underline{n}$, and whose corresponding eigenvalue, μ, is on the unit circle. Suppose $P_k = P_e + U_k$ where $U_0 = \beta_0\underline{v}\,\underline{v}^*$. Then an analysis similar to the one used in the proof of Theorem 3.23 shows that $U_k = \beta_k\underline{v}\,\underline{v}^*$ where β_k is described by the scalar difference equation

$$\beta_{k+1} = \frac{\beta_k}{1 + \mu^*(\underline{v}^*D\underline{v})\beta_k} \quad . \tag{3.196}$$

The factor $\mu^*(\underline{v}^*D\underline{v})$ in the denominator of (3.196) is equal to

$$\mu^*(\underline{v}^*D\underline{v}) = \underline{v}^*(M'V^{-1}MP_e + I)^{-1}M'V^{-1}M\underline{v}$$

$$= \underline{v}^*M'(MP_eM' + V)^{-1}M\underline{v} \tag{3.197}$$

which implies that it is positive since if $M\underline{v}$ were equal to $\underline{0}$, then $(B-P_eD)\underline{v} = \Phi\underline{v} = \mu\underline{v}$ and the mode would be unobservable. The stability of (3.196) can be determined from an examination of Figure 3.4, which shows that β_k converges toward zero when

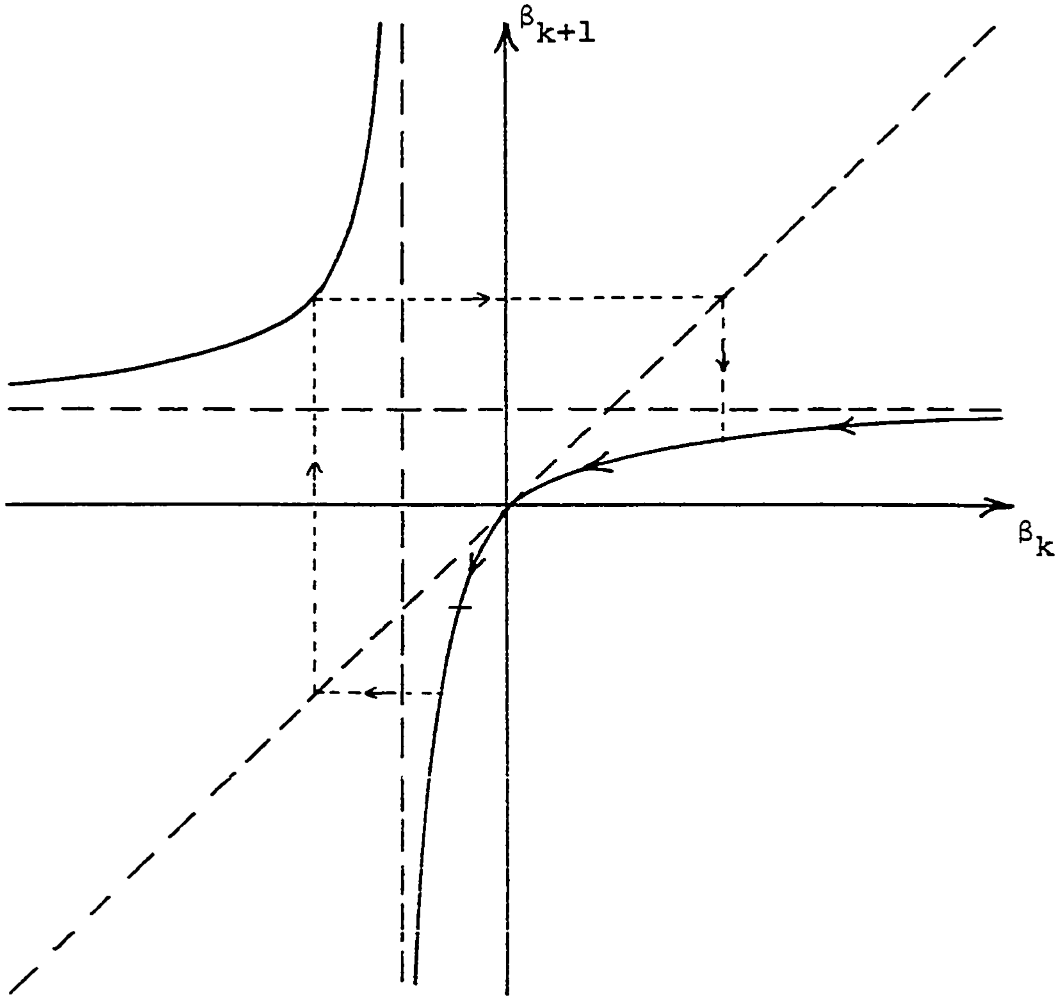


Figure 3.4. Graph of β_{k+1} versus β_k

its initial value is positive but it diverges from zero when its initial value is even slightly negative. Thus since $\underline{n}^* P_k \underline{n} = \beta_k (\underline{v}^* \underline{n})^2$, it can be seen that there are indefinite P_k matrices arbitrarily near P_e which diverge from P_e as stated in conclusion (a).

Next suppose the system has one or more unstable modes which are undriven and observable. Let the columns of N_1 and N_2 be basis vectors of η_1 and η_2, the subspaces of $\eta(R_d)$ associated with stable and unstable modes respectively, let n_1 and n_2 be the dimensions of η_1 and η_2, and let the columns of L be basis vectors of $\eta_2^{\perp}$. Then the following relationships are true:

$$H[N_1 \quad N_2] = 0 \ , \tag{3.198}$$

$$\Phi'[N_1 \quad N_2] = [N_1 \quad N_2] \operatorname{diag}(\Phi_1, \Phi_2) \tag{3.199}$$

where $|\lambda_i(\Phi_1)| < 1$ and $|\lambda_i(\Phi_2)| > 1$,

$$\Phi L = L\Phi_{\ell}, \tag{3.200}$$

$$H = LGL^* \text{ where } G \geq 0, \tag{3.201}$$

$$R_d = LR_g L^* \tag{3.202}$$

where

$$R_g = \sum_{i=0}^{m-1} \Phi_{\ell}^{\,i} G (\Phi_{\ell}^{\,i})^K \tag{3.203}$$

which is the state covariance matrix for the reduced system

$$\underline{z}_{k+1} = \Phi_{\ell} \underline{z}_k + \underline{g}_k \tag{3.204}$$

$$\underline{y}_k = ML\underline{z}_k + \Delta\underline{y}_k \tag{3.205}$$

where $E[\underline{g}_k\underline{g}_k^*] = G$. Thus if $P_0N_2 = 0$, then $P_k = LX_kL^*$ where X_k is the solution of the reduced covariance equation

$$X_{k+1} = \Phi_\ell[X_k - X_kL^*M'(MLX_kL^*M' + V)^{-1}MLX_k]\Phi_\ell^* + G \ . \tag{3.206}$$

Now $(L^*N_1)\Phi_1 = L^*\Phi'N_1 = \Phi_\ell^*(L^*N_1)$ and $G(L^*N_1) = 0$ so the n_1 columns of L^*N_1 are linearly independent null vectors of R_g, and since

$$\text{nullity}(R_g) + n_2 = \text{nullity}(R_d) = n_1 + n_2 \ , \tag{3.207}$$

these columns are basis vectors for $\eta(R_g)$. Thus since the eigenvalues of Φ_1 are within the unit circle, the only undriven modes in the reduced system are stable modes. Therefore (3.206) has only one nonnegative definite equilibrium solution, X_e, and any $X_0 \geq 0$ causes $X_k \to X_e$ and $P_k \to LX_eL^* = P_{e2}$. Thus the first part of conclusion (b) is proved.

The author has investigated two other approaches toward a complete analytic proof of conclusions (a) and (b). The results of this investigation are stated in the remainder of this section. Although these results do not completely achieve the stated objective, they do show that further investigation of these approaches is warranted. Probably the most promising approach is a study of the properties of the following solution of (3.150):

$$U_k = (B-P_eD)^k U_0[(DP_e-C)^k + S_kU_0]^{-1} \tag{3.208}$$

where

$$S_k = (DP_e-C)^{k-1}D + (DP_e-C)^{k-2}D(B-P_eD) + \dots$$
$$+ (DP_e-C)D(B-P_eD)^{k-2}+D(B-P_eD)^{k-1}$$
$$= (DP_e-C)^kE-E(B-P_eD)^k \tag{3.209}$$

and where E satisfies the equations

$$(DP_e-C)E-E(B-P_eD) = D \tag{3.210}$$

$$(DP_e-C)E(DP_e-C)^* - E = (\Phi')^{-1}(M'V^{-1}M+M'V^{-1}MP_eM'V^{-1}M)\Phi^{-1}. \tag{3.211}$$

The fact that a closed-form solution of (3.150) can be written is somewhat surprising since the usual object of a stability analysis is to determine the large time behavior of a differential or difference equation whose solution is not known. But in spite of the simple form of (3.208), its properties are not readily apparent due to the fact that an inverse of a sum is involved. However, if (3.208) is re-written in the form

$$U_k = (B-P_eD)^kU_0[I+EU_0-(B^*-D^*P_e)^kE(B-P_eD)^kU_0]^{-1}(B^*-D^*P_e)^k \tag{3.212}$$

and if the eigenvalues of $B-P_eD$ are within the unit circle and $I + EU_0$ is nonsingular, then it is apparent that $U_k \to 0$.

The conditions necessary for the eigenvalues of $B-P_eD$ to be within the unit circle are known from Section 3: P_e must be nonnegative definite, no random walk mode may be either undriven or unobservable, and no unstable mode may be unobservable. However, the conditions which result in a nonsingular $I + EU_0$ matrix are not known, and the resolution of this problem would be an interesting and useful topic for further investigation.

Lyapunov stability theory provides another promising approach to the problem. Let

$$\underline{\mu}(t_k) = \begin{bmatrix} \underline{u}_1(t_k) \\ \underline{u}_2(t_k) \\ \cdot \\ \cdot \\ \cdot \\ \underline{u}_n(t_k) \end{bmatrix} \tag{3.213}$$

where $\underline{u}_i(t_k)$ is the $i\underline{th}$ column of U_k. Then (3.150) is equivalent to the equation

$$\underline{\mu}(t_{k+1}) = F(\underline{\mu}(t_k))\underline{\mu}(t_k) \tag{3.214}$$

where $F(\underline{\mu}(t_k))$ is the Kronecker product[1]

$$F(\underline{\mu}(t_k)) = (B-P_eD) \times [(DP_e-C) + DU_k]^{-1} . \tag{3.215}$$

A possible Lyapunov function for (3.214) is

---

[1]See Bellman [25] for the definition and properties of the Kronecker product of two matrices.

$$V = \underline{\mu}^*(t_k)S\underline{\mu}(t_k) \tag{3.216}$$

where S is a symmetric, positive definite matrix equal to the solution of the equation

$$S-F^*(\underline{0})SF(\underline{0}) = I, \tag{3.217}$$

and where $F(\underline{0}) = (B-P_eD) \times (B-P_eD)^*$. Numerical examples indicate that (3.216) has all of the properties of a Lyapunov function provided either all of the eigenvalues of $B-P_eD$ are inside the unit circle and no element of η_2 is a null vector of P_k, or if an element of η_2 is a null vector of P_k, then $B-P_eD$ must have an eigenvalue equal to λ_d, the eigenvalue of an undriven unstable mode, with the remainder of the eigenvalues inside the unit circle.

The function (3.216) is not a valid Lyapunov function when the random process has a random walk mode which is undriven and observable because $F(\underline{0})$ has an eigenvalue on the unit circle, and therefore no solution exists for S in (3.217). This difficulty is partially avoided by defining the Lyapunov function as

$$V = \underline{\mu}^*(t_k)S(t_k)\underline{\mu}(t_k), \tag{3.218}$$

where $S(t_k)$ is the solution of the equation

$$S(t_k) - F^*(\underline{\mu}(t_k))S(t_k)F(\underline{\mu}(t_k)) = I\ . \tag{3.219}$$

This function exists and, by numerical examples, has the properties of a Lyapunov function, provided the random walk mode is simple and U_k has no null vector in ν, the subspace of $\eta(R_d)$ associated with this mode. If U_k does have such a null vector, then $F(\underline{\mu}(t_k))$ has an eigenvector on the unit circle, so again no solution for $S(t_k)$ exists. This problem can possibly be overcome by the change of coordinates $P_k = LX_kL^*$ where the columns of L are basis vectors for $\nu^{\perp}$. If the random walk mode is not simple, then (3.218) fails to be a Lyapunov function because it is not continuous at $\underline{\mu}(t_k) = \underline{0}$. However, numerical examples indicate that the first difference, ΔV, of (3.218) is negative along any trajectory of U_k for which $U_0 \geq 0$, which possibly indicates that a simple modification of (3.218) may prove to be a valid Lyapunov function for multiple-dimensional, as well as simple, random walk modes.

The following example illustrates these comments and shows the rather complex behavior that the solutions of the covariance equation have when the random process has a multiple-dimensional random walk mode which is undriven and observable.

<u>Example 3.2</u>:

The random process has a 2-dimensional random walk mode which is undriven and observable. The Φ matrix and its Jordan decomposition, $\Phi = T\Delta T^{-1}$, are given by the equation

$$\Phi = \begin{bmatrix} -1 & 4/3 \\ -3 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix}^{-1} , \quad (3.220a)$$

and the H, M, and V matrices are equal to

$$H = 0, \quad (3.220b)$$

$$M = [-1 \quad 1], \quad (3.220c)$$

$$V = 1 \; . \quad (3.220d)$$

The only equilibrium solution of the covariance equation in this case is $P_e = 0$. Therefore $B-P_eD = B = \Phi$, $DP_e-C = -C = (\Phi')^{-1}$, and (3.150) takes the form

$$U_{k+1} = \Phi U_k[I + M'V^{-1}MU_k]^{-1}\Phi. \quad (3.221)$$

When U_k is small, the solution of (3.221) is nearly identical to the solution of the linearized equation $U_{k+1} = \Phi U_k \Phi'$ which is unstable since Φ has a multiple eigenvalue on the unit circle. However, the nonlinearity in (3.221) causes those solutions whose initial values are nonnegative definite to eventually approach 0. This is most apparent when the Jordan expansion $\Phi = T\Lambda T^{-1}$ is used to replace (3.221) by the pair of equations

$$U_k = TX_kT \quad (3.222)$$

$$X_{k+1} = \Lambda X_k[I + T'M'V^{-1}MTX_k]^{-1}\Lambda', \quad (3.223)$$

which in expanded form are equal to

$$U_k = x_{11}\begin{bmatrix}4 & 6\\ 6 & 9\end{bmatrix} + x_{12}\begin{bmatrix}-2 & 0\\ -3 & 0\end{bmatrix} + x_{21}\begin{bmatrix}-2 & -3\\ 0 & 0\end{bmatrix} + x_{22}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} \tag{3.224}$$

and

$$X_{k+1} = \begin{bmatrix} \dfrac{x_{11}+x_{12}+x_{21}+x_{22}}{1+x_{11}+x_{12}+x_{21}+x_{22}} & \dfrac{x_{12}+x_{22}}{1+x_{11}+x_{12}+x_{21}+x_{22}} \\ \dfrac{x_{21}+x_{22}}{1+x_{11}+x_{12}+x_{21}+x_{22}} & \dfrac{x_{22}+(x_{11}x_{22}-x_{12}x_{21})}{1+x_{11}+x_{12}+x_{21}+x_{22}} \end{bmatrix} \tag{3.225}$$

where x_{ij} is the (i, j)th element of X_k. Note that $U_k \geq 0$ if and only if the conditions $x_{11} \geq 0$, $x_{22} \geq 0$, $x_{11} + x_{12} + x_{21} + x_{22} \geq 0$, and $x_{11}x_{22} - x_{12}x_{21} \geq 0$ are all satisfied.

Now as a representative example, consider the initial condition

$$X_0 = \begin{bmatrix}0 & 0\\ 0 & a\end{bmatrix} \tag{3.226}$$

where $0 < a << 1$. The general solution of (3.225) is then

$$x_{11}(t_k) = \frac{ak^2}{1 + a\sum_{i=1}^{k} i^2} \tag{3.227}$$

$$x_{12}(t_k) = x_{21}(t_k) = \frac{ak}{1 + a \sum\limits_{i=1}^{k} i^2} \qquad (3.228)$$

$$x_{22}(t_k) = \frac{a}{1 + a \sum\limits_{i=1}^{k} i^2} \qquad (3.229)$$

which is plotted in Figure 3.5. Note that when X_k is small, the system is essentially unstable because at each iteration the relatively constant value of x_{22} is added to x_{12} and x_{21} which in turn are added to x_{11}. But when the x_{11}/x_{22} ratio becomes sufficiently large, the effect of the divisor $1 + x_{11} + x_{12} + x_{21} + x_{22}$ predominates and the elements of X_k begin to decrease in value. On the other hand when a slightly negative, the divisor accelerates the divergence of X_k from 0.

The behavior of the function (3.218) resulting from two different solutions of the covariance equation are plotted in Figure 3.6. In both curves, the U_0 matrix was chosen such that X_0 had the form of (3.226) with a = .01 for curve (a) and with a = .1 for curve (b). It is interesting to observe that although the elements of both X_k and U_k increase in value for the first 5 or 6 samples of either solution, the value of the function (3.218) decreases monotonically. However, the initial value of V increases as the parameter a decreases. This means that

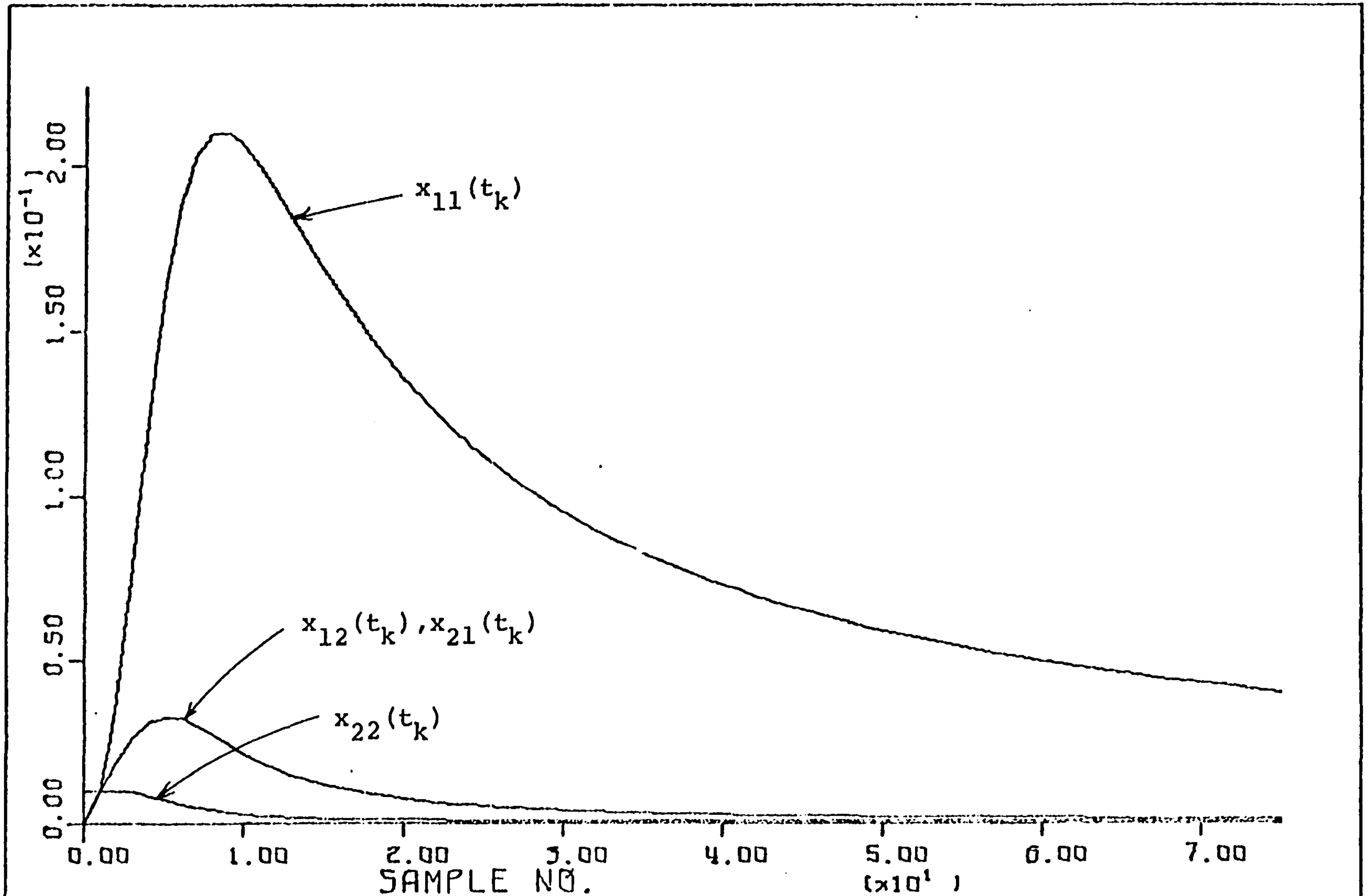


Figure 3.5. Graph of the transient behavior of the variables $x_{ij}(t_k)$ for Example 3.2 with a=0.01

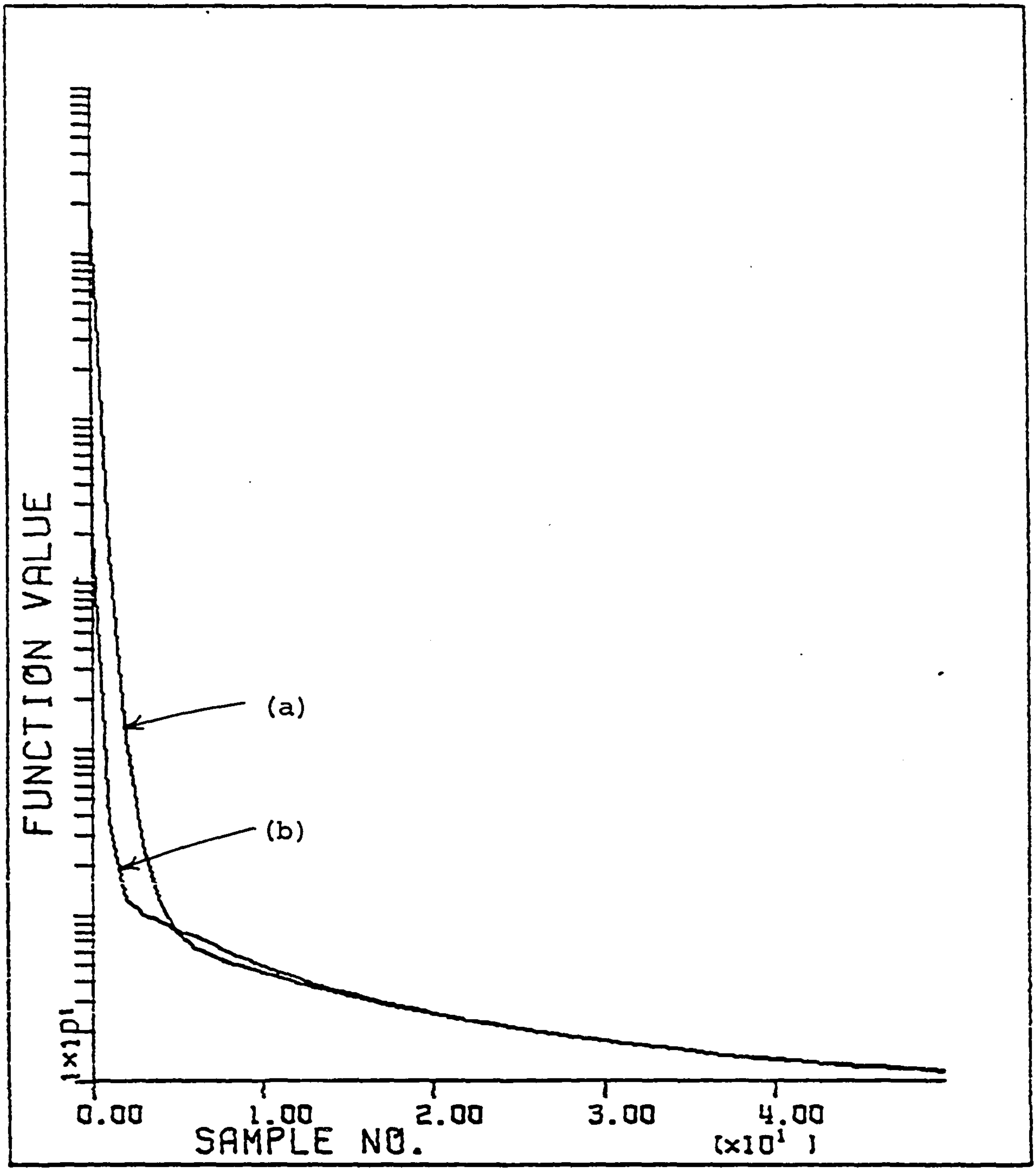


Figure 3.6. Graphs of function (3.218) versus k along two different trajectories of U_k

Curve (a) -- a = 0.01

Curve (b) -- a = 0.1

$$\lim_{U_k \to 0} V \neq 0 \tag{3.230}$$

for all paths of approach, as it must if (3.218) is to be a valid Lyapunov function.

Another interesting aspect of this case is the effect of round-off error in the computer calculation of P_k. Since the system is nearly unstable for small P_k, it is not surprising that round-off errors have a significant effect on the computed P_k matrices. In fact if round-off error should happen to make the x_{22} component of U_k negative, the solution would be expected to diverge toward more negative definite values. To prevent this from happening, Sorenson [26, p. 261] proposes using the equation

$$P_{k+1} = \Phi[(I - K_k M)P_k(I - K_k M)' + K_k V K_k']\Phi' + H \tag{3.231}$$

where

$$K_k = P_k M'(MP_k M' + V)^{-1} \tag{3.232}$$

rather than any other form of the covariance equation because every term in (3.231) is nonnegative definite. However when round-off error is present, this is no guarantee that the computed P_{k+1} matrix will be nonnegative definite since a considerable amount of cancellation can occur in the computation of the first two terms of (3.231). This is demonstrated in Figure 3.7 in which the trace of the computed P_k matrix is plotted vs. k when P_k is computed in three

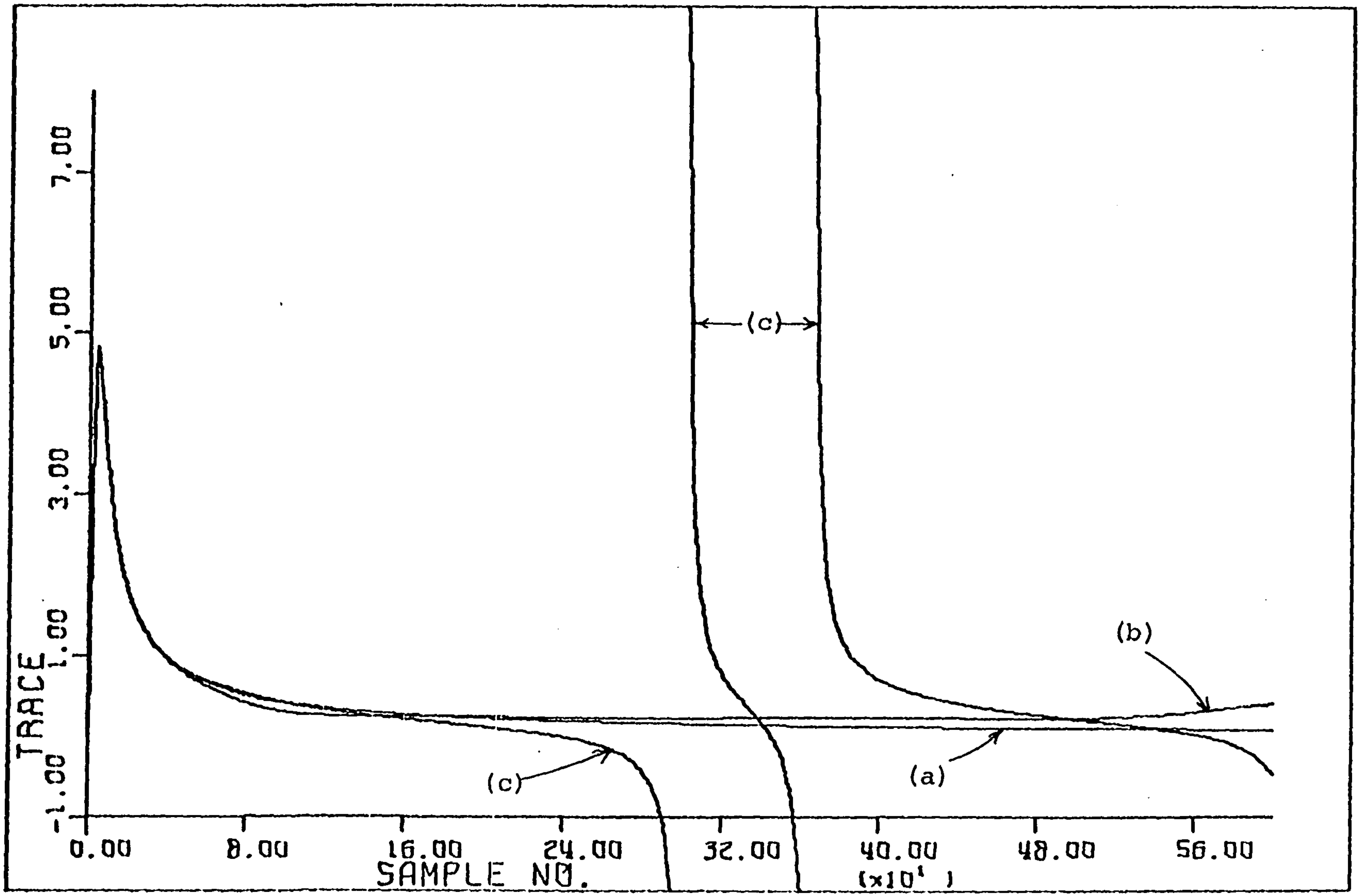


Figure 3.7. Transient behavior of $tr(P_k)$ as computed by three different algorithms ((a) by (3.233)using double precision arithmetic, (b) by (3.233) using single precision arithmetic, (c) by (3.234) using single precision arithmetic

computationally different, but algebraically equivalent, ways using (3.231). For curves (a) and (b), the elements of P_{k+1} were computed according to the formula

$$p_{ij}(t_{k+1}) = f_{i1}p_{11}f_{j1} + f_{i1}p_{12}f_{j2} + f_{i2}p_{21}f_{j1} + f_{i2}p_{22}f_{j2} + g_iVg_j + h_{ij} , \qquad (3.233)$$

where $f_{i\ell}$, g_i, h_{ij}, and $p_{\ell m}$ are the elements of $\Phi(I-K_kM)$, ΦK_k, H, and P_k respectively, using double precision floating point arithmetic on the IBM 360/65 for curve (a) and using single precision floating point arithmetic for curve (b). For curve (c), the elements of P_{k+1} were computed according to the formula

$$p_{ij}(t_{k+1}) = f_{i1}(p_{11}f_{j1}+p_{12}f_{j2}) + f_{i2}(p_{21}f_{j1}+p_{22}f_{j2}) + g_iVg_j + h_{ij} \qquad (3.234)$$

using single precision floating point arithmetic. The initial P matrix was

$$P_0 = \begin{bmatrix} .1 & 0 \\ 0 & 0 \end{bmatrix} \qquad (3.235)$$

in all cases. The author has not made a detailed study of the manner in which round-off errors enter into the computation of (3.233, 3.234), but evidently the factoring of f_{i1} and f_{i2} in (3.234) is more prone to negative round-off errors.

c. Stability of the actual covariance matrix In many applications of the Kalman filter, it is quite difficult to evaluate the actual driving, measurement and initial error covariance matrices, H_a, V_a, and $P_a(t_0)$ respectively. This means that H, V, and P_0 are only estimates of the corresponding actual covariance matrices, and therefore the computed error covariance matrix, P_k, may not be equal to the actual error covariance, $P_a(t_k)$. In fact it will be shown that $P_a(t_k)$ can be unbounded even though P_k is bounded.

The difference equation for $P_a(t_k)$ is derived as follows. The actual random process is described by the equations

$$\underline{x}_{k+1} = \Phi\underline{x}_k + \underline{h}_k \tag{3.236}$$

$$\underline{y}_k = M\underline{x}_k + \underline{\Delta y}_k \tag{3.237}$$

where $E[\underline{h}_k\underline{h}_k']=H_a$ and $E[\underline{\Delta y}_k\underline{\Delta y}_k']=V_a$, and the Kalman filter estimate of the state vector is given by the equation

$$\hat{\underline{x}}_{k+1} = \Phi[\hat{\underline{x}}_k + K_k(\underline{y}_k - M\hat{\underline{x}}_k)] \tag{3.238}$$

where

$$K_k = P_kM'(MP_kM'+V)^{-1} \tag{3.239}$$

and

$$P_{k+1} = \Phi[(I-K_kM)P_k(I-K_kM)'+K_kVK_k']\Phi'+H\ . \tag{3.240}$$

Therefore the difference equations for the actual error,

$\underline{e}_k = \hat{\underline{x}}_k - \underline{x}_k$, and the actual error covariance are

$$\underline{e}_{k+1} = \Phi[(I - K_k M)\underline{e}_k + K_k \underline{\Delta v}_k] - \underline{h}_k \tag{3.241}$$

and

$$P_a(t_{k+1}) = \Phi[(I- K_k M)P_a(t_k)(I - K_k M)' + K_k V_a K_k']\Phi' + H_a \tag{3.242}$$

respectively. Notice that (3.240, 3.242) have the same form, but (3.242) is a time-varying linear equation while (3.240) is effectively nonlinear since K_k depends on P_k.

To begin the stability analysis of (3.242), let it be assumed that P_k is equal to one of the equilibrium solutions, P_e, of (3.240). Then K_k becomes the constant matrix

$$K_e = P_e M'(MP_e M' + V)^{-1} , \tag{3.243}$$

and by (1.28f, 3.50b, 3.50d)

$$\Phi(I - K_e M) = B-P_e D ,[1] \tag{3.244}$$

so (3.242) becomes the time-invariant difference equation

$$P_a(t_{k+1}) = (B-P_e D)P_a(t_k)(B-P_e D)^* + \Phi K_e V_a K_e^* \Phi' + H_a . \tag{3.245}$$

The stability of (3.245) depends on the location of the eigenvalues in the Jordan form, K, of $B-P_e D$. If the system has no random walk or unstable modes which are undriven,

---

[1] In the time-varying case, these same equations imply the identity $\Phi(I-K_k M) = B-P_{k+1}D$.

then the eigenvalues of $B-P_eD$ are within the unit circle and $P_a(t_k)$ approaches the equilibrium value P_{ae} which is the solution of the linear equation

$$P_{ae} - (B-P_eD)P_{ae}(B-P_eD)^* = \Phi K_e V_a K_e^* \Phi' + H_a \ . \qquad (3.246)$$

Furthermore if $H_a = H$ and $V_a = V$, then $P_{ae} = P_e$, so $P_a(t_k) \rightarrow P_e$ even though the initial covariances $P_a(t_0)$ and P_0, may not have been equal. Thus in this case it is possible to be rather careless in the selection of P_0.

If the random process has a simple random walk mode which is undriven, then there exists a vector, $\underline{n}$, such that $\Phi'\underline{n} = \lambda^*\underline{n}$ with $|\lambda| = 1$ and such that $P_e\underline{n} = H\underline{n} = \underline{0}$. Therefore there also exists a vector, $\underline{v}$, such that $(B-P_eD)\underline{v} = \lambda\underline{v}$ and $\underline{n}^*\underline{v} = 1$, while $\underline{n}$ is orthogonal to all other eigenvectors of $B-P_eD$. Thus $\alpha_k\underline{v}\ \underline{v}^*$ is one of the components in the expansion of $P_a(t_k)$ in terms of the eigenvectors of $B-P_eD$, and by pre- and postmultiplication of (3.245) by $\underline{n}^*$ and $\underline{n}$, the difference equation for α_k is

$$\alpha_{k+1} = \alpha_k + \underline{n}^* H_a\underline{n} \ . \qquad (3.247)$$

This implies that α_k either increases without bound if the random walk mode is not actually undriven, or it remains fixed at a value which is greater than zero[1] if the actual

---

[1]One of the components in the expansion of P_k is $\beta_k\underline{v}\ \underline{v}^*$ where β_k, which is described by the nonlinear difference equation (3.196), approaches zero if its initial value is nonnegative.

variance of this mode is not zero when P_k first becomes equal to P_e. In both cases, the actual estimation errors could be considerably larger than the computed error covariance matrix would indicate.

If the random process has an unstable mode which is undriven, then there exists a vector, $\underline{n}$, such that $\Phi'\underline{n} = \lambda^*\underline{n}$ with $|\lambda| > 1$ and such that $H\underline{n} = \underline{0}$, and if P_0 is positive semidefinite with $P_0\underline{n} = \underline{0}$, then $P_k\underline{n} = \underline{0}$ for all $k \geq 0$ and P_k approaches a positive semidefinite equilibrium matrix for which $P_e\underline{n} = \underline{0}$. Therefore, there also exists a sequence of eigenvectors, $\underline{v}_k$, such that $\Phi(I - K_kM)\underline{v}_k = \lambda\underline{v}_k$ with $\underline{n}^*\underline{v}_k = 1$, while $\underline{n}$ is orthogonal to all other eigenvectors of $\Phi(I - K_kM)$. Therefore $\alpha_k\underline{v}_k\underline{v}_k^*$ is one of the components in the expansion of $P_a(t_k)$, and by pre- and postmultiplication of (3.242) by $\underline{n}^*$ and $\underline{n}$, the difference equation for α_k is

$$\alpha_{k+1} = |\lambda|^2\alpha_k + \underline{n}^*H_a\underline{n} \ . \tag{3.248}$$

Since $P_k \to P_e$, the eigenvectors $\underline{v}_k \to \underline{v}$ where $\underline{v}$ is a vector such that $(B-P_eD)\underline{v} = \lambda\underline{v}$ with $\underline{n}^*\underline{v} = 1$, so the component $\alpha_k\underline{v}_k\underline{v}_k^*$ of $P_a(t_k)$ remains bounded if and only if α_k is bounded. Therefore by (3.248), $P_a(t_k)$ is unbounded if the unstable mode is not actually undriven or if the initial variance in the estimate of this mode is not actually zero as it was assumed to be when P_0 was chosen such that $P_0\underline{n} = \underline{0}$.

Now suppose that P_k is not equal to one of the equilibrium matrices. By Theorem 3.24, it is known that if all of the random walk and unstable modes of the random process are both driven and observable, then P_k approaches the equilibrium matrix, P_e, whose corresponding K matrix has all of its eigenvalues within the unit circle. And by the preceding analysis of (3.245), it is known that when $P_k = P_e$, these same conditions imply that $P_a(t_k)$ approaches P_{ae}. Therefore it is expected that $P_a(t_k)$ should approach P_{ae} for any pair of nonnegative definite initial error covariances, P_0 and $P_a(t_0)$, and this is shown to be true by the following theorem.

Theorem 3.25:

If all unstable and random walk modes of the random process are both driven and observable and if P_0 and $P_a(t_0)$ are symmetric nonnegative definite, then $P_a(t_k) \to P_{ae}$ as $k \to \infty$.

Proof:

Consider first the vector difference equation

$$\underline{x}_{k+1} = \Phi(I-K_k M)\underline{x}_k \tag{3.249}$$

whose solution can be written as

$$\underline{x}_m = \Phi_x(t_m, t_0)\underline{x}_0 \tag{3.250}$$

where

$$\Phi_x(t_m,t_0) = \Phi(I-K_{m-1}M) \ldots \Phi(I - K_0 M) \ . \tag{3.251}$$

Let P_e denote the symmetric, nonnegative definite equilibrium solution of the covariance equation, and let the function $V(\underline{x}_k)$ be defined as

$$V(\underline{x}_k) = \underline{x}_k{}' S \underline{x}_k \tag{3.252}$$

where S is the solution of the equation

$$S - (I-K_e M)' \Phi' S \Phi (I-K_e M) = I \ . \tag{3.253}$$

This solution is symmetric, positive definite since the eigenvalues of $\Phi(I-K_e M)$ are within the unit circle. The first difference of (3.252) along any trajectory of (3.249) is

$$\Delta V = V(\underline{x}_{k+1}) - V(\underline{x}_k)$$

$$= \underline{x}_k{}' [(I-K_k M)' \Phi' S \Phi (I-K_k M) - S] \underline{x}_k \ . \tag{3.254}$$

Since the model of the random process is regular, $P_k \rightarrow P_e$ and $K_k \rightarrow K_e$ which implies that there exists an integer, N, such that the quantity within brackets in (3.254) is negative definite for all $k \geq N$. This implies that (3.252) is a Lyapunov function for (3.249) and that (3.249) is asymptotically stable in the large.

Now let $P_a(t_k) = P_{ae} + W_k$. Then W_k is described by the difference equation

$$W_{k+1} = \Phi(I-K_k M) W_k (I-K_k M)' \Phi' + D_k \tag{3.255}$$

where

$$D_k = \Phi(I-K_kM)P_{ae}(I-K_kM)' - (I-K_eM)P_{ae}(I-K_eM)'$$

$$+ K_kV_aK_k' - K_eV_aK_e']\Phi', \qquad (3.256)$$

and the solution of (3.255) for W_m is

$$W_m = \Phi_x(t_m,t_0)W_0\Phi_x'(t_m,t_0)$$

$$+ \sum_{i=0}^{m-1} \Phi_x(t_m,t_{i+1})D_i\Phi_x'(t_m,t_{i+1}) \; . \qquad (3.257)$$

Since (3.249) is asymptotically stable in the large, there exist positive constants c_1 and λ_1 such that

$$||\Phi_x(t_m,t_j)|| \leq c_1e^{-\lambda_1(m-j)} \; , \qquad (3.258)$$

and since $P_k \to P_e$ exponentially, there exist positive constants c_2 and λ_2 such that

$$||D_i|| \leq c_2e^{-\lambda_2 i} \; . \qquad (3.259)$$

Therefore

$$||W_m|| \leq c_1{}^2e^{-2\lambda_1 m} + \sum_{i=0}^{m-1} c_1{}^2c_2e^{-2\lambda_1(m-i-1)-\lambda_2 i} \qquad (3.260)$$

which approaches 0 as $m\to\infty$. Thus $P_a(t_k) \to P_{ae}$. Q.E.D.

As was the case in the analysis of the computed covariance equation, a relaxation of the completely driven condition on random walk or unstable modes leads to much more interesting results. These are illustrated by the following numerical examples.

Example 3.3:

The random process has a stable mode which is driven and observable and an unstable mode which is undriven and observable. The Φ matrix and its Jordan decomposition, $\Phi T = T\Lambda$, are given by

$$\begin{bmatrix} 3/2 & -2/3 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 3/2 \end{bmatrix} \qquad (3.261a)$$

the H, H_a, and M matrices are equal to

$$H = H_a = \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix} \qquad (3.261b)$$

$$M = [-1 \quad 1] \ , \qquad (3.261c)$$

and $V = V_a = 1$. The covariance equation has two nonnegative definite equilibrium solutions:

$$P_{e1} = \begin{bmatrix} 13.13719 & 10.61281 \\ 10.61281 & 11.88719 \end{bmatrix} \qquad (3.262)$$

which is positive definite and stable, and

$$P_{e2} = \begin{bmatrix} 4.53113 & 6.79669 \\ 6.79669 & 10.19504 \end{bmatrix} \qquad (3.263)$$

which is positive semidefinite and unstable. The vector $\underline{n} = [3 \quad -2]'$ is a null vector of P_{e2} and H and is an eigenvector of Φ' associated with the unstable eigenvalue. The traces of P_k and $P_a(t_k)$ are plotted in Figure 3.8 for the initial conditions

$$P_0 = \begin{bmatrix} 64 & 0 \\ 0 & 64 \end{bmatrix} \quad P_a(t_0) = \begin{bmatrix} 49 & 0 \\ 0 & 49 \end{bmatrix} \tag{3.264a,b}$$

and in Figure 3.7 for the initial conditions

$$P_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad P_a(t_0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{3.265a,b}$$

Both curves in Figure 3.8 approach the trace of P_{e1} because $\underline{n}$ is not a null vector of the first P_0 matrix. This implies that $P_k \to P_{e1}$, and since $B-P_{e1}D$ is stable and $H_a = H$ and $V_a = V$, $P_a(t_k)$ also approaches P_{e1}. However in Figure 3.9, the trace of P_k initially approaches the trace of P_{e2} and the trace of $P_a(t_k)$ increases without bound because $\underline{n}$ is a null vector of the second P_0 matrix. And this would continue indefinitely if perfect computation were possible, but since P_{e2} is unstable and since some round-off error is present, even with double precision arithmetic, the numerical value for P_k eventually departs from P_{e2} and approaches P_{e1}. When this happens, the difference equation for $P_a(t_k)$ becomes stable, and therefore the trace of $P_a(t_k)$ decreases rapidly

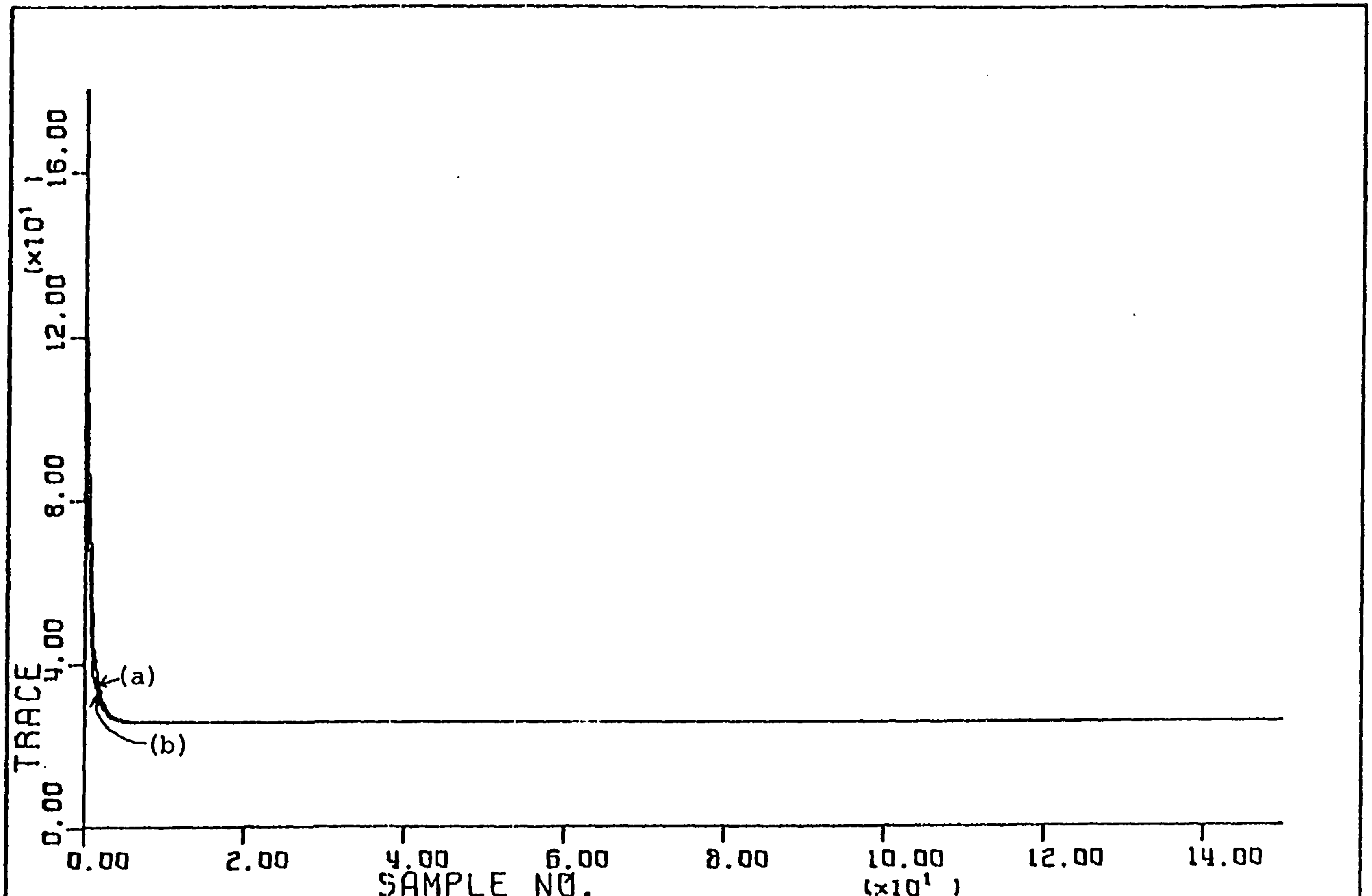


Figure 3.8. Transient behavior of the trace of the computed and actual covariance matrices for Example 3.3 with $P_0 = 64I$ and $P_a(t_0)=49I$ (Curve (a) - trace of P_k, Curve (b) - trace of $P_a(t_k)$)

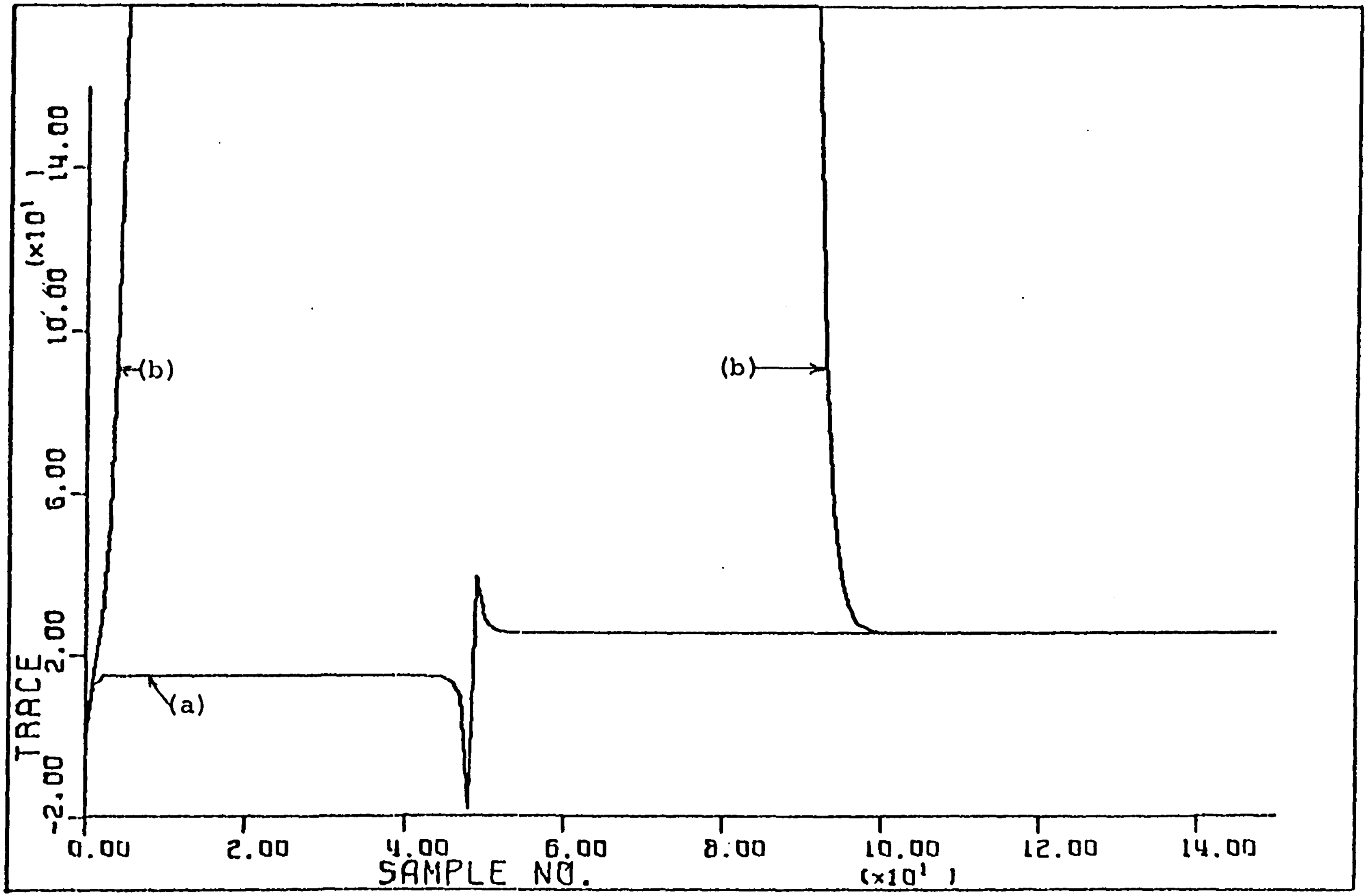


Figure 3.9. Transient behavior of the trace of the computed and actual covariance matrices for Example 3.3 with $P_0 = 0$ and $P_0(t_0)=I$ (Curve (a) - trace of P_k, Curve (b) -- trace of $P_a(t_0)$)

to $tr(P_{e1})$ from the large value which accumulated during the time when P_k was equal to P_{e2}.

Example 3.4:

The random process has a stable mode which is driven and observable and a random walk mode which is undriven and observable. The Φ matrix and its Jordan decomposition, $\Phi T = T\Lambda$, are given by

$$\begin{bmatrix} 1 & -1/3 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad (3.266)$$

and the H, H_a, M, V, and V_a matrices have the same values as in Example 3.3. There is only one nonnegative definite P_e matrix in this case, which happens to be equal to P_{e2} of the previous example. The $B-P_eD$ matrix has, as expected, a simple eigenvalue on the unit circle, and since $\underline{n} = [3 \quad -2]$ is a null vector of H_a as well as P_e and H, (3.247) implies that the component $\alpha_k \underline{v}\, \underline{v}^*$ of $P_a(t_k)$ remains equal to the value it has when P_k first becomes essentially equal to P_e. Figures 3.10 and 3.11 indicate that the final value of α_k depends on the initial covariance, P_0 and $P_a(t_0)$. The traces of P_k and $P_a(t_k)$ are plotted in Figure 3.10 for the initial conditions (3.264) and in Figure 3.11 for the initial conditions (3.265). These figures demonstrate that when P_0 is considerably greater than P_e, then $P_a(t_k)$ rapidly

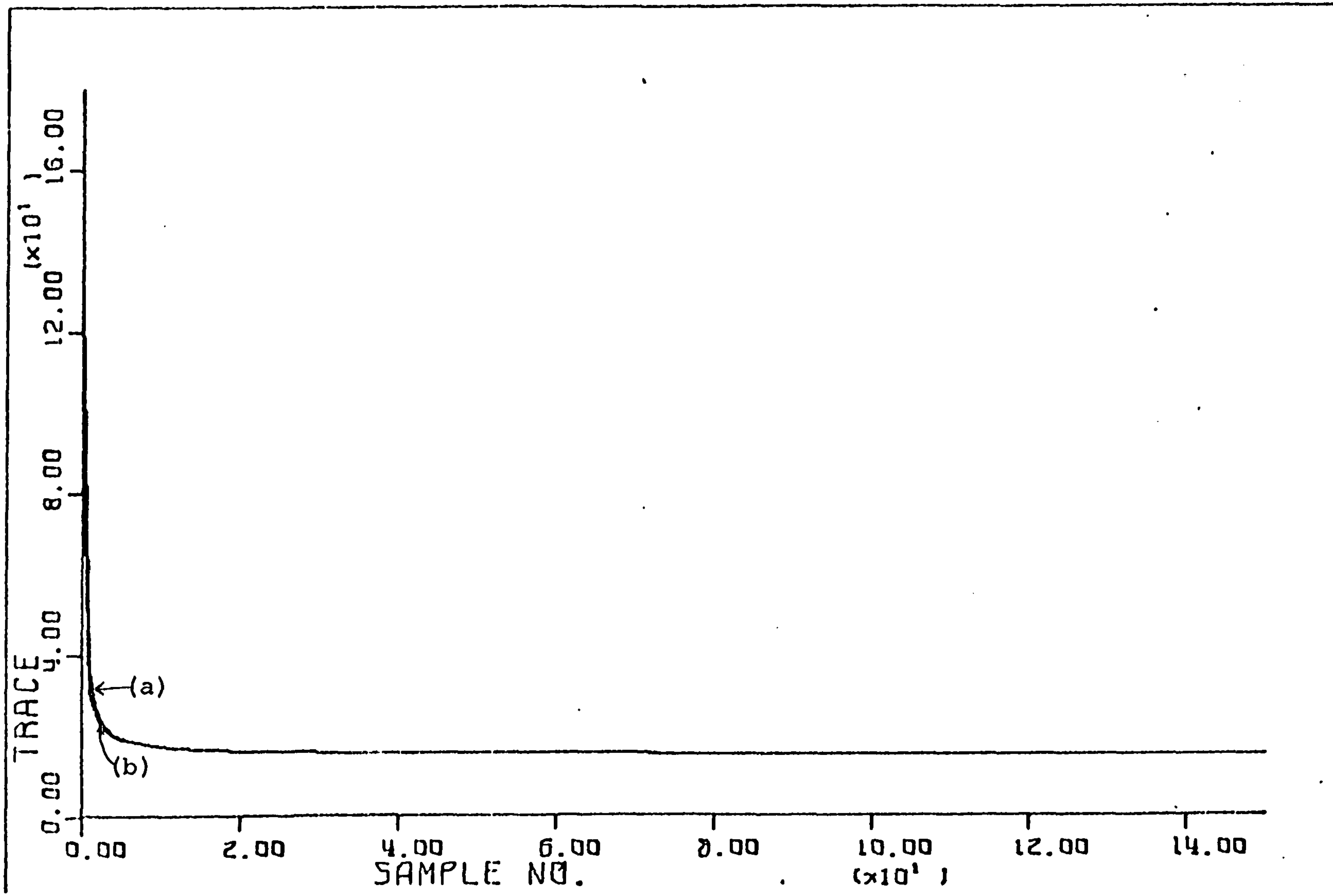


Figure 3.10. Transient behavior of the trace of the computed and actual covariance matrices for Example 3.4 with $P_0=64I$ and $P_a(t_0)=49I$ (Curve (a) - trace of P_k, Curve (b) - trace of $P_a(t_k)$)

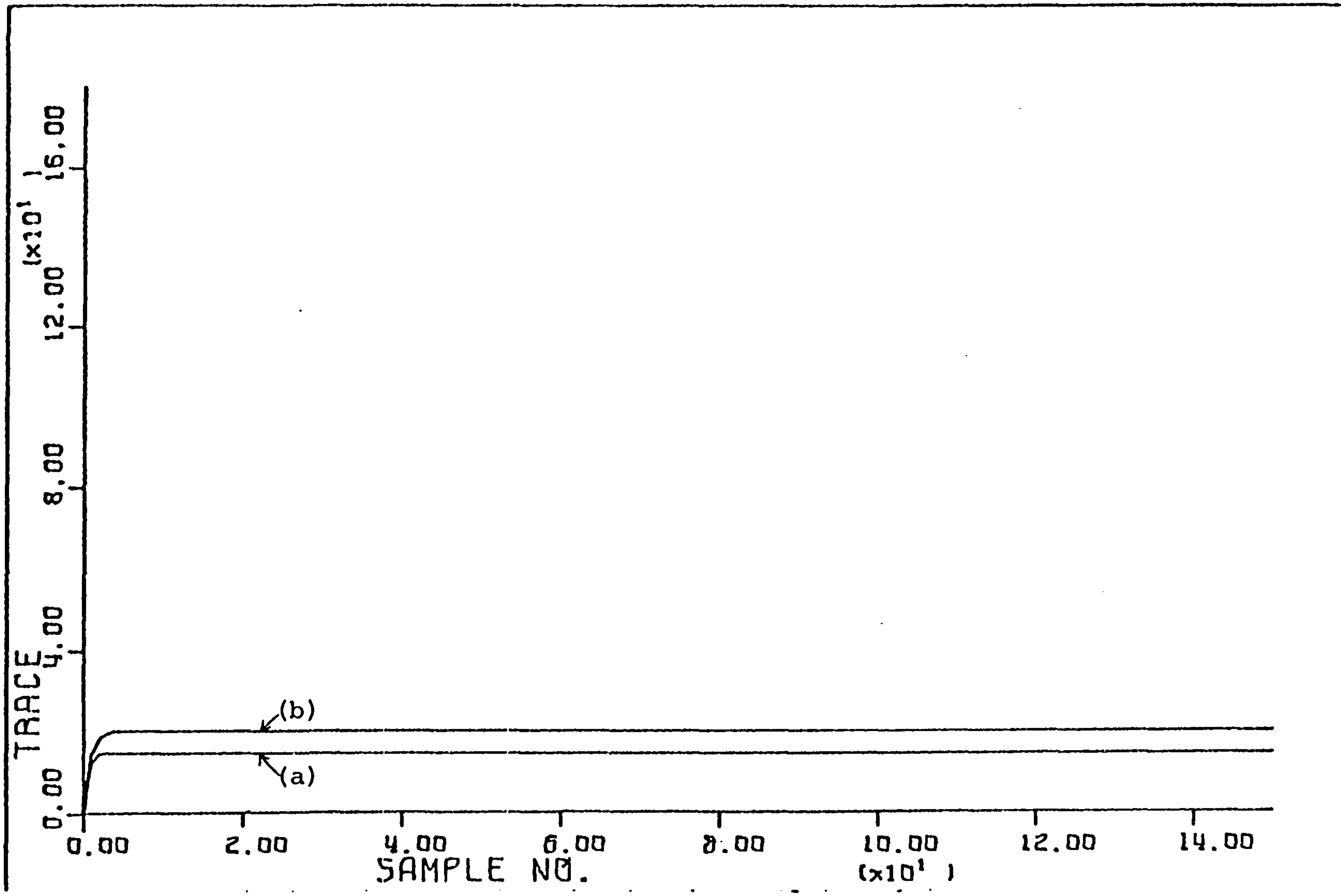


Figure 3.11. Transient behavior of the trace of the computed and actual covariance matrices for Example 3.4 with $P_0=0$ and $P_a(t_0)=I$ (Curve (a) - trace of P_k, Curve (b) - trace of $P_a(t_k)$)

approaches P_k and both of them approach P_e, but when $\underline{n}$ is a null vector of P_0 and not a null vector of $P_a(t_0)$, then the final value of $P_a(t_k)$ is significantly greater than the final value of P_k.

Example 3.5:

The random process has a multiple random walk mode which is undriven. The Φ, H, M, and V matrices are given by (3.220), and $V_a = 1$. The only equilibrium solution in this case is $P_e = 0$, and the corresponding $B-P_eD$ matrix has a multiple eigenvalue on the unit circle. Equation (3.245) is therefore unstable, and it is expected that (3.242) may also be unstable when P_k is near zero. This is verified by Figures 3.12 and 3.13 in which the traces of P_k and $P_a(t_k)$ are plotted for the initial conditions

$$P_0 = \begin{bmatrix} .1 & 0 \\ 0 & 0 \end{bmatrix} \qquad P_a(t_0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.267a,b}$$

and

$$P_0 = \begin{bmatrix} 40 & 60 \\ 60 & 90 \end{bmatrix} \qquad P_a(t_0) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{3.268a,b}$$

respectively with $H_a = 0$, and by Figure 3.14 in which the initial conditions are

$$P_0 = \begin{bmatrix} .01 & 0 \\ 0 & 0 \end{bmatrix} \qquad P_a(t_0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.269a,b}$$

with

$$H_a = \begin{bmatrix} 2 \times 10^{-6} & 0 \\ 0 & 0 \end{bmatrix}. \tag{3.270}$$

Figures 3.12 and 3.13 indicate that the boundedness of $P_a(t_k)$ depends on the initial conditions. In the first case both P_k and $P_a(t_k)$ converge to P_e, but in the second case, P_k converges to P_e while $P_a(t_k)$ diverges. The reason for this divergence can best be seen by making the change of variables: $\Phi = T\Lambda T^{-1}$, $P_k = TX_kT^*$, and $P_a(t_k) = TY_kT^*$. Then by (3.224, 3.225)

$$X_k = \begin{bmatrix} x_{11}(t_k) & 0 \\ 0 & 0 \end{bmatrix} \tag{3.271}$$

where

$$x_{11}(t_{k+1}) = \frac{x_{11}(t_k)}{1 + x_{11}(t_k)}, \tag{3.272}$$

and by (3.242)

$$y_{k+1} = \begin{bmatrix} \dfrac{y_{11}+y_{12}+y_{21}+y_{22}+x_{11}^2}{(1+x_{11})^2} & \dfrac{y_{12}+y_{22}}{1+x_{11}} \\ \dfrac{y_{21}+y_{22}}{1+x_{11}} & y_{22} \end{bmatrix} \tag{3.273}$$

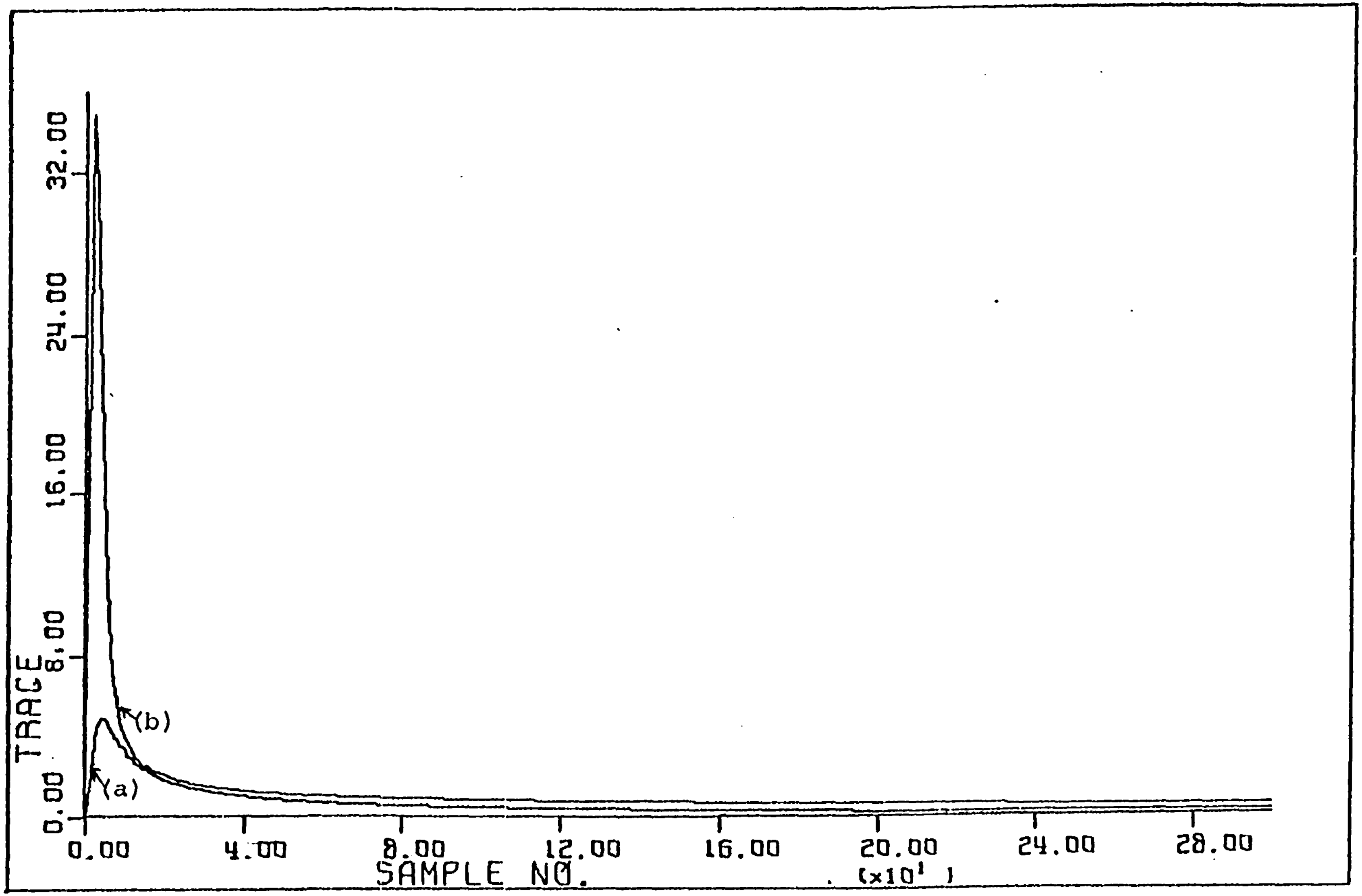


Figure 3.12. Transient behavior of the trace of the computed and actual covariance matrices for Example 3.5 with $H_a=0$ and the first set of initial conditions (Curve (a) - trace of P_k^a, Curve (b) - trace of $P_a(t_k)$)

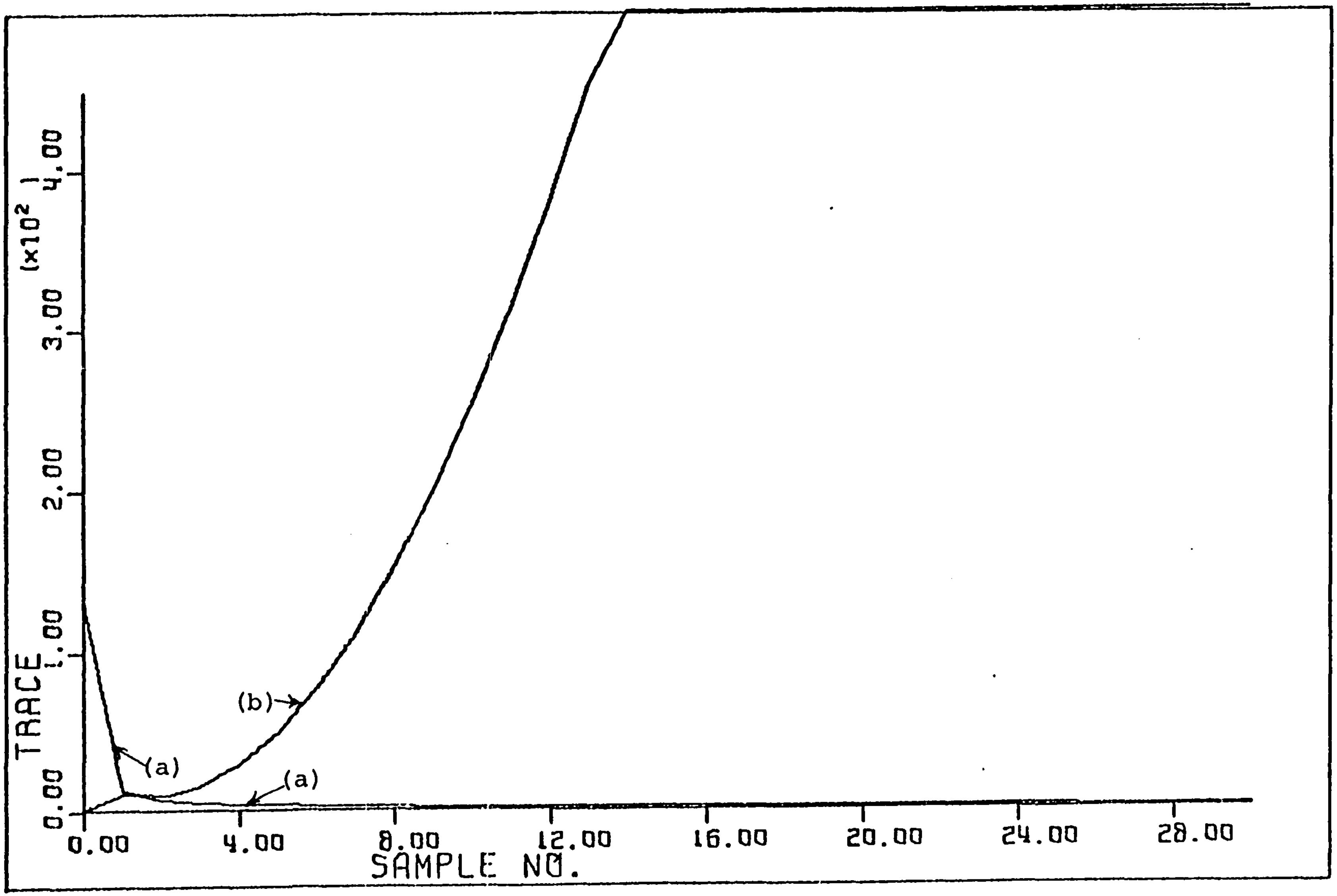


Figure 3.13. Transient behavior of the trace of the computed and actual covariance matrices for Example 3.5 with $H_a=0$ and the second set of initial conditions (Curve (a) - trace of P_k^a, Curve (b) - trace of $P_a(t_k)$)

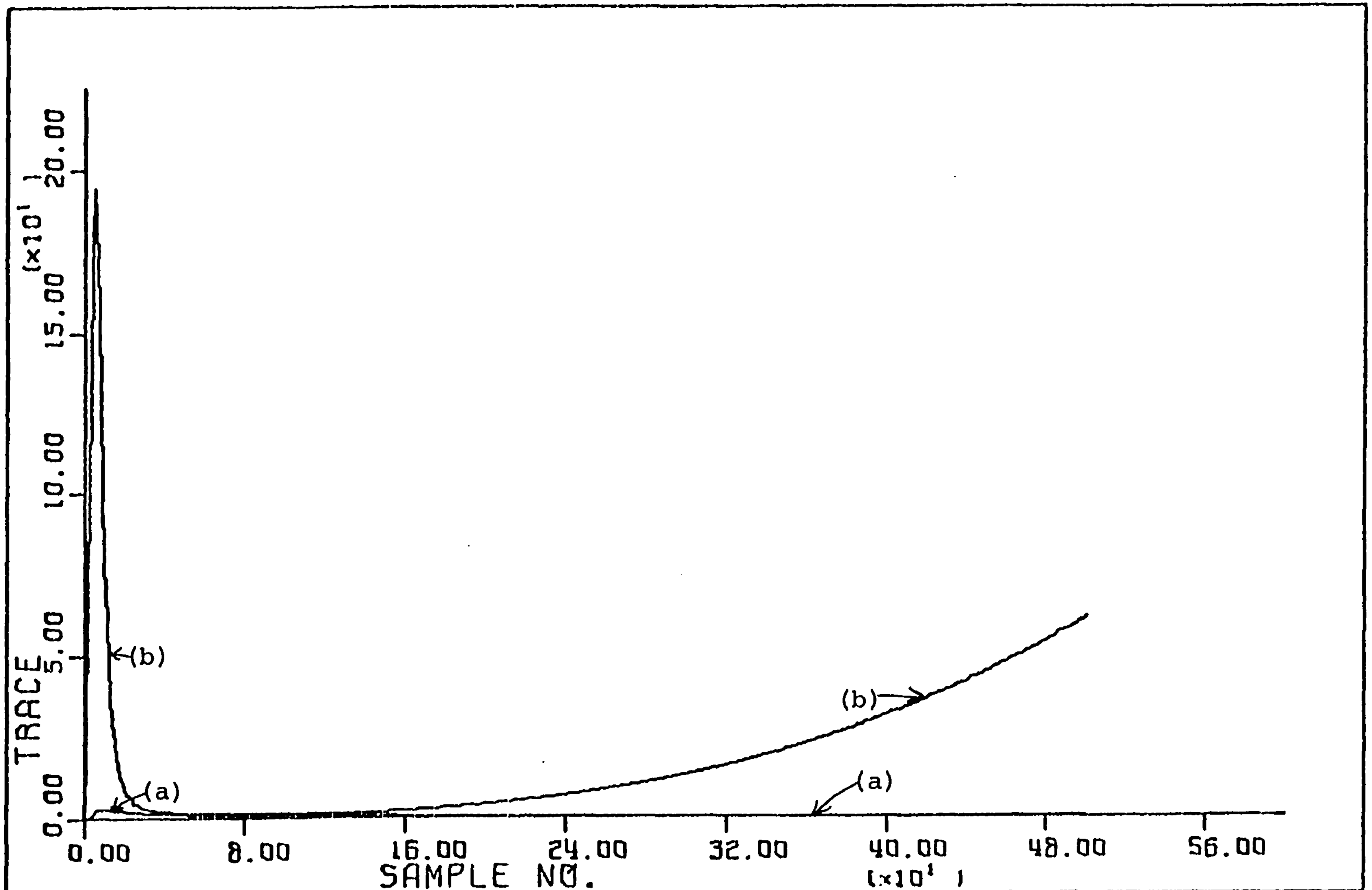


Figure 3.14. Transient behavior of the trade of the computed and actual covariance for Example 3.5 with $H_a \geq 0$ and the third set of initial conditions (Curve (a) - trace of $\bar{P}_k$, Curve (b) - trace of $P_a(t_k)$)

where x_{11} and y_{ij} denote $x_{11}(t_k)$ and $y_{ij}(t_k)$ respectively. This equation implies that Y_k, and therefore $P_a(t_k)$, is unbounded unless $P_a(t_0)$ has the form

$$P_a(t_0) = y_{11} \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}, \tag{3.274}$$

and since (3.269b) is not of this form, $P_a(t_k)$ diverges. Figure 3.14 shows that if the actual random process is slightly driven, then the actual covariance matrix eventually diverges from the computed covariance matrix, even though they may have apparently been equal for a significant period of time.

These examples show that when a random walk or unstable mode is modeled as being undriven, the boundedness of $P_a(t_k)$ depends in a complex manner on the initial covariances, P_0 and $P_a(t_0)$, as well as on whether the mode is actually undriven. An analytic description of the set of initial covariances which result in a bounded solution of the difference equation for $P_a(t_k)$ would be an interesting and useful topic for further research.

IV. SUMMARY OF RESULTS AND CONCLUSIONS

In this chapter, the main results of the analysis of Chapter III are summarized, the conclusions resulting from this analysis are stated, and some possibilities for future research are discussed.

A. Summary of Results

The following list summarizes the main results of the previous chapter.

1. It has been shown that the discrete-time and continuous-time covariance equations have analogous existence, symmetry, definiteness, and ordering properties.

2. A graphical analysis of the scalar, time-invariant covariance equation clearly shows how the stability of this equation is affected by the stability of the random process and by whether it is driven and/or observable.

3. Potter's [15] procedure for solving the quadratic matrix equation $A + BP + PB^{*} - PDP = 0$ has been extended to the solution of the more general equation $A + BP + PC - PDP = 0$. It has been shown that all of the solutions can still be expressed as $P = F\Gamma^{-1}$ where F and Γ are the upper and lower halves of a solution of the equation $RT = TJ$, and some conditions which must be satisfied by the Jordan matrix, J, have been derived.

4. It has been shown that the equilibrium solutions of the time-invariant covariance equation satisfy a quadratic matrix equation. The R matrix used in the solution of this equation is called R_p, and its eigenvalues and eigenvectors have some special properties which have been stated and proved.

5. It has been shown that the existence, symmetry, definiteness, and local stability properties of an equilibrium solution can be deduced from the eigenvalues of the corresponding J matrix.

6. The relationships between the symmetry and definiteness properties of the a posteriori and a priori equilibrium matrices have been derived. Also a quadratic matrix equation for the a posteriori equilibrium matrices was derived and some relationships between its solution and the solution of the quadratic matrix equation for the a priori equilibrium matrices were shown.

7. It has been shown that the computed covariance matrix is globally stable when its initial value is nonnegative definite and the random process is regular, i.e. it has neither random walk nor unstable modes which are either undriven or unobservable.

8. The effects of having an undriven random walk or unstable mode in the random process, and the effect that the initial covariance matrix has in such cases, has been

discussed. Although this discussion is only partially based on analytic proofs, these proofs and some numerical examples give a fairly clear picture of the global stability properties of the covariance equation under these conditions.

9. It has been shown that the stability of the difference equation for the actual, as opposed to the computed, covariance matrix depends upon the model that is assumed for the random process. If the assumed model is regular, then this difference equation is asymptotically stable in the large, but if the assumed model is not regular, then its stability depends upon the initial values of both the actual and computed covariance matrices as well as upon the accuracy of the assumed model.

B. Conclusions

In this section, the main conclusions resulting from the analysis of Chapter III are listed. These conclusions supply answers to all four of the questions stated at the beginning of Chapter II.

1. If the random process model is regular, then the difference equation for the computed covariance matrix, P_k, has a single nonnegative definite equilibrium solution, P_e, and any solution whose initial value is nonnegative definite approaches P_e as $k\to\infty$. Furthermore the difference equation for the actual covariance matrix, $P_a(t_k)$, is a stable, time-

varying, linear difference equation whose solution approaches the equilibrium value P_{ae} of (3.101). If the assumed model is not an accurate representation of the actual random process, then P_k and $P_a(t_k)$ follow different trajectories and converge to different equilibriums. They both remain bounded, however, and there is only a gradual degradation in the performance of the filter as the model departs from the actual random process. Also since every solution whose initial value is nonnegative definite converges to the same equilibrium value, small errors in the numerical solution of the covariance equation have a relatively minor effect on the operation of the filter. This is even true when P_e is singular and P_k is slightly indefinite due to round-off error. Since P_e is locally stable, P_k returns to P_e for any initial condition within a suitably small region of P_e.

2. The covariance equation is unstable if there is an unstable mode which is unobservable or if there is a random walk mode which is driven and unobservable. If there is a random walk mode which is both undriven and unobservable, then the covariance equation has a continuum of equilibrium solutions resulting from the variance in the estimate of the random walk mode which is constant and equal to the arbitrary initial variance.

3. If the model of the random process has a random walk mode which is undriven and observable, then the difference

equation for the computed covariance matrix, P_k, has a single nonnegative definite equilibrium solution, P_e, which is singular and only semi-stable, i.e. every solution whose initial value is nonnegative definite converges toward P_e, but there are indefinite P_0 matrices arbitrarily near P_e which cause P_k to diverge from P_e. Therefore in this case, small errors in the numerical solution of the covariance equation can have catastrophic effects on the operation of the filter if these errors happen to generate an indefinite P_k matrix which is in the set of solutions which diverge from P_e. This cannot happen until the filter has operated long enough to reduce the variance in the estimate of the random walk mode to a value which is roughly equal to the magnitude of the round-off errors, and even then it is difficult to predict whether it will happen. Nevertheless, the possibility of this divergence implies that this model is not a very "safe" model, and it should be used with great care.

Another problem with this model is that the transition matrix in the difference equation for $P_a(t_k)$ has an eigenvalue on the unit circle when P_k is equal to P_e. This causes $P_a(t_k)$ to be unbounded if the random walk mode is not actually undriven, and it allows $P_a(t_k)$ to approach an equilibrium value considerably different from P_e if the mode is actually undriven.

4. If the model of the random process has an unstable mode which is undriven and observable, then the difference equation for the computed covariance matrix has two non-negative definite equilibrium solutions, P_{e1} which is positive definite and stable and P_{e2} which is positive semi-definite and unstable. The only solutions which approach P_{e2} are those that start from an initial covariance matrix for which $P_0\underline{n} = \underline{0}$, where $\underline{n}$ is a null vector of P_{e2} and H and is an eigenvector of Φ' associated with the unstable eigenvalue. These solutions cause the difference equation for the actual covariance matrix to be unstable and its solutions remain bounded only if the unstable mode is actually undriven and $P_a(t_0)\underline{n}$ is exactly $\underline{0}$. Therefore for this model, initial conditions such that $P_0\underline{n} = \underline{0}$ are also not very safe and should be used very carefully. All other nonnegative definite P_0 matrices, however, cause $P_k \to P_{e1}$ and $P_a(t_k) \to P_{ae}$, both of which are stable equilibriums. Small computational errors have a significant effect only on those solutions that approach the unstable equilibrium, P_{e2}. Since it is unstable, these errors are propagated forward with increasing magnitude, and eventually cause P_k to depart from P_{e2} and settle down to P_{e1}. However as Figure 3.9 shows, the end result of this can be beneficial since the difference equation for $P_a(t_k)$ then becomes stable.

C. Future Research Possibilities

In this section, some possible topics for further research are discussed. First, the topics which involve the strengthening or completion of theorems which have been proposed on the basis of observed numerical phenomena are listed.

1. Can the conditions of Theorems 3.13 through 3.16 be shown to be both necessary and sufficient? In each case, the theorems, as stated, prove only one side of the double implication which numerical examples seem to indicate.

2. Can conditions be found which are both necessary and sufficient for the R_p matrix to be derogatory? And when R_p is derogatory, is there a way to describe the resulting continuum of equilibrium solutions in terms of the least number of arbitrary parameters? Also can Theorems 3.13 through 3.17 be extended to include these cases as well?

3. Can it be proved that every solution, whose initial value is nonnegative definite, approaches a unique nonnegative definite equilibrium matrix when the random process has a random walk mode which is undriven and observable? Likewise when the random process has an unstable mode which is undriven and observable, can it be proved that the solutions of the covariance equation whose initial values are nonnegative definite approach either P_{e_1} or P_{e_2} as described in item 3 of the previous section? In particular, are the

solutions for which $P_0\underline{n} = \underline{0}$ the only solutions which approach P_{e_2}? These conclusions, which are presently based upon numerical examples, might possibly be proved either by further analysis of the expressions (3.208) or (3.212) or by finding a valid Lyapunov function.

The remaining topics represent areas of research that are supplemental to, rather than improvements of, the research reported in this thesis.

4. It was noted in the discussion of Example 3.5 that the convergence of the actual covariance matrix, $P_a(t_k)$, to an equilibrium value depends upon the initial values of both the computed, P_0, and actual, $P_a(t_0)$, covariance matrices when the model of the random process has either a random walk or an unstable mode which is undriven and observable. An analytic description of the set of $P_a(t_0)$ matrices which, for a given value of P_0, cause $P_a(t_k)$ to converge to its equilibrium value, P_{ae}, would be very useful.

5. It has been shown in this thesis that the covariance equation has a stable, nonnegative definite equilibrium solution provided all of the random walk and unstable modes in the random process are completely observable. Recall from Chapter I that a mode, $\underline{\xi}_i(t_k)$, is defined to be completely observable if its subspace, τ_i, is contained within the range of the observability matrix, P_d. This definition, therefore, provides only a yes/no answer concerning the

observability of these modes, rather than a numerical measure of their observability. This implies that the question of stability of the covariance equation is likewise answered with an absolute yes or no. It is proposed that a relative measure of stability may be obtained by a modification of the definition of relative observability proposed by R. G. Brown [27].

6. It has been assumed throughout most of this thesis that the measurement error covariance matrix, V, is positive definite, and in fact, V^{-1} appears in many of the formulas. An investigation of what happens when V becomes positive semidefinite would be quite useful.

7. Finally, a very useful topic would be an investigation of the time-varying case to see whether the results that have been derived in this thesis for the time-invariant case also apply when the Φ_k, M_k, V_k, and H_k matrices are functions of k.

V. LITERATURE CITED

1. Kalman, R. E. A new approach to linear filtering and prediction problems. ASME Transactions, Series D, 82: 35-45. 1960.

2. Schlee, F. H., C. J. Standish, and N. F. Toda. Divergence in the Kalman filter. AIAA Journal 5: 1114-1120. 1967.

3. Toda, N. F., F. H. Schlee, and P. Obsharsky. The region of Kalman filter convergence for several autonomous navigation modes. AIAA Paper 67-623. August, 1967.

4. Kalman, R. E. and R. S. Bucy. New results in linear filtering and prediction theory. ASME Transactions, Series D, 83: 95-108. 1961.

5. Zadeh, L. A. and C. A. Desoer. Linear system theory. New York, N.Y., McGraw-Hill. 1963.

6. Bryson, A. E. and D. E. Johansen. Linear filtering for time-varying systems using measurements containing colored noise. IEEE Transactions on Automatic Control AC-10: 4-10. 1965.

7. Ho, Y. C. The method of least squares and optimal filtering theory. Research Memorandum RM-3329-PR, Rand Corp. 1962.

8. Kalman, R. E. Contributions to the theory of optimal control. Sociedad Mathematica Mexicana 5: 102-120. 1960.

9. Kalman, R. E. On the general theory of control systems. International Congress on Automatic Control, Moscow, Proceedings 1: 481-493. 1960.

10. Kalman, R. E. and J. E. Bertram. Control system analysis and design via the 'second method' of Lyapunov, Continuous-time systems. ASME Transactions, Series D, 82: 371-393. 1960.

11. Kalman, R. E. and J. E. Bertram. Control system analysis and design via the 'second method' of Lyapunov, Discrete-time- systems. ASME Transactions, Series D, 82: 394-400. 1960.

12. Kalman, R. E. New methods in Weiner filtering theory. In Bogdanoff, J. L. and F. Kozin, eds. Engineering Applications of Random Function Theory and Probability Symposium. Pp. 270-308. New York, N.Y., John Wiley and Sons. 1963.

13. Potter, J. E. A matrix equation arising in statistical filtering theory. National Aeronautics and Space Administration Report CR-270. August, 1965.

14. Nishimura, T. Error bounds of continuous Kalman filters and the application to orbit determination problems. IEEE Transactions on Automatic Control AC-12: 268-275. 1967.

15. Potter, J. E. Matrix quadratic solutions. Journal Applied Math. 14: 496-501. 1966.

16. Nishimura, T. On the a priori information in sequential estimation problems. IEEE Transactions on Automatic Control AC-11: 197-204. 1966.

17. Nishimura, T. Correction to and extension of 'On the a priori information in sequential estimation problems'. IEEE Transactions on Automatic Control AC-12: 123. 1967.

18. Heffes, H. The effect of erroneous models on the Kalman filter response. IEEE Transactions on Automatic Control AC-11: 541-543. 1966.

19. Deyst, J. J. and C. F. Price. Conditions for asymptotic stability of the discrete minimum-variance linear estimator. IEEE Transactions on Automatic Control AC-13: 702-705. 1968.

20. Price, C. F. An analysis of the divergence problem in the Kalman filter. IEEE Transactions on Automatic Control AC-13: 699-702. 1968.

21. Penrose, R. A generalized inverse for matrices, Part 3. Cambridge Philosophical Society Proceedings 51: 406-413. 1955.

22. Penrose, R. On best approximate solutions of linear matrix equations, Part 1. Cambridge Philosophical Society Proceedings 52: 17-19. 1956.

23. Bellman, R. Introduction to matrix analysis. New York, N.Y., McGraw-Hill Book Co. 1959.

24. Gantmacher, F. R. The theory of matrices. Vols. 1 and 2. New York, N.Y., Chelsea Publishing Co. 1960.

25. Bellman, R. E. Kronecker products and the second method of Lyapunov. Mathematishe Nachrichten 20: 17-19. 1959.

26. Sorenson, H. W. Kalman filtering techniques. In Leondes, C. T., ed. Advances in control systems. Vol. 3. Pp. 219-292. New York, N.Y., Academic Press. 1966.

27. Brown, R. G. Not just observable, but how observable? National Electronics Conference Proceedings 22: 409-414. 1966.

VI. ACKNOWLEDGMENTS

The author is indebted to Dr. R. G. Brown for suggesting this problem and for the guidance given during its investigation. This research was partially supported by the Office of Naval Research, contract number NOOO14-68-A-0162.

VII. APPENDIX A: NOTATIONAL CONVENTIONS

In this appendix, the notational conventions that have been used in this thesis are listed.

1. Equation numbers are enclosed in parentheses and reference numbers are enclosed in brackets.

2. Scalars are denoted by lower case letters, matrices are denoted by upper case letters, and column vectors are denoted by lower case underlined letters. A particular component of a matrix or column vector is denoted by a subscripted lower case letter, e.g. a_{ij} denotes the (i,j)th element of the matrix A. The identity matrix is denoted by I, the null matrix by 0, and the null vector by $\underline{0}$.

3. The transpose of a vector or matrix is denoted by the prime symbol, and the complex conjugate transpose is denoted by the asterisk.

4. Vector spaces are generally denoted by lower case Greek letters, but the range space of a matrix, say A, is denoted by $\mathcal{R}(A)$ and its null space is denoted by $\eta(A)$.

5. The determinant of a matrix, A, is denoted by $|A|$, while $||A||$ and $||\underline{a}||$ denote the norm of the matrix A and the vector $\underline{a}$ respectively. The methods used in computing the norm of a matrix or vector are specified in those portions of the text where the norms are used. The ith eigenvalue of A is denoted by $\lambda_i(A)$.

6. The expected value of a random variable, say $\underline{x}$, is denoted by $E[\underline{x}]$, and the derivative of a variable with respect to time is denoted by $\dot{\underline{x}}$.

VIII. APPENDIX B: SOME MATRIX THEOREMS

Some theorems concerning matrix theory are stated and proved in this appendix. The first theorem is a corollary of the matrix inversion lemma.

Theorem 8.1:

If $MP_kM' + V$ is nonsingular, then the a posteriori covariance matrix, Q_e, is related to the a priori covariance matrix, P_e, by the equations

$$Q_k = P_k(I + M'V^{-1}MP_k)^{-1} \tag{8.1a}$$

$$= (I + P_kM'V^{-1}M)^{-1}P_k \; . \tag{8.1b}$$

Proof:

Theorem 3.20 implies that the right side of (8.1a) exists. The equality of this expression with Q_k is shown by the following sequence of identities, starting from (1.24):

$$Q_k = P_k - P_kM'(MP_kM'+V)^{-1}MP_k$$

$$= P_k[I-M'(MP_kM'+V)^{-1}MP_k](I+M'V^{-1}MP_k)(I+M'V^{-1}MP_k)^{-1}$$

$$= P_k[I+M'V^{-1}MP_k-M'(MP_kM'+V)^{-1}MP_k(I+M'V^{-1}MP_k)]$$

$$\cdot (I+M'V^{-1}MP_k)^{-1}$$

$$= P_k[I+M'V^{-1}MP_k-M'(MP_kM'+V)^{-1}(MP_kM'+V)V^{-1}MP_k]$$

$$\cdot \; (I+M'V^{-1}MP_k)^{-1}$$

$$= P_k(I+M'V^{-1}MP_k)^{-1} \ . \tag{8.2}$$

Identity (8.1b) is proved in a similar manner. Q.E.D.

Notice that the nonsingularity of P_k is not required in the statement or proof of these identities. This, therefore is the primary advantage of these identities over the matrix inversion lemma, (1.28d).

The next theorem establishes a similarity relation that is used in the proof of Theorem 3.10.

Theorem 8.2:

Let Ω denote the simple Jordan matrix $\omega I + N$, where N supplies the superdiagonal ones. Then $(\Omega^*)^{-1}$ is similar to the matrix $(\omega^*)^{-1}I + N$.

Proof:

Let n denote the number of rows and columns in Ω, and let X be the matrix whose elements are equal to

$$\begin{aligned} x_{ij} &= (-1)^{j-1}\begin{pmatrix} j-1 \\ i-n \end{pmatrix}(\omega^*)^{n+j-i-1} && \text{when } j > n-i \\ &= 0 && \text{when } j \le n-i \ . \end{aligned} \tag{8.3}$$

For example, when n=4, X has the form

$$X = \begin{bmatrix} 0 & 0 & 0 & -\mu^6 \\ 0 & 0 & \mu^4 & -\mu^5 \\ 0 & -\mu^2 & \mu^3 & -\mu^4 \\ 1 & -\mu & \mu^2 & -\mu^3 \end{bmatrix} \tag{8.4}$$

where $\mu = \omega^*$. By direct substitution, it can be verified that X satisfies the equation

$$X = \Omega^* X[(\omega^*)^{-1}I + N].$$

Therefore X also satisfies the equation

$$(\Omega^*)^{-1}X = X[(\omega^*)^{-1}I + N],$$

which implies the conclusion since X is nonsingular. When Ω is the direct sum of a set of simple Jordan matrices, the same conclusion is obtained by defining X to be the direct sum of a set of matrices of the form (8.3). Q.E.D.

The third theorem establishes a property of cyclic subspaces which is used in the proof of Theorem 3.18. A subspace, σ, of an n-dimensional vector space is called cyclic if it is spanned by a sequence of vectors of the form

$$\underline{x},\ A\underline{x},\ \ldots,\ A^{p-1}\underline{x} \tag{8.7}$$

where A is an n-by-n matrix, $\underline{x}$ is an arbitrary element of the vector space, and p is the smallest integer such that the vector $A^p\underline{x}$ is a linear combination,

$$A^p\underline{x} = (a_1A^{p-1} + a_2A^{p-2} + \ldots + a_pI)\underline{x} \,, \tag{8.8}$$

of the preceding p vectors of the sequence. Equation (8.8) can be written as $\phi(A)\underline{x} = \underline{0}$ where $\phi(\lambda)$ is the monic polynomial

$$\phi(\lambda) = \lambda^p - a_1\lambda^{p-1} - \ldots - a_p \,, \tag{8.9}$$

which is called the minimal annihilating polynomial of $\underline{x}$.

Theorem 8.3:

If σ is a cyclic subspace associated with the matrix A, then σ contains an eigenvector of A.

Proof:

Let λ_i be a zero of $\phi(\lambda)$, let

$$\theta(\lambda) = \frac{\phi(\lambda)}{\lambda-\lambda_i} \,, \tag{8.10}$$

and let

$$\underline{v} = \theta(\lambda)\underline{x} \,. \tag{8.11}$$

The vector $\underline{v}$ is not zero because the degree of $\theta(\lambda)$ is less than the degree of the minimal annihilating polynomial, $\phi(\lambda)$. Also

$$(A-\lambda_iI)\underline{v} = (A-\lambda_iI)[\theta(A)\underline{x}] = \phi(A)\underline{x} = \underline{0}.$$

Thus $\underline{v}$ is an eigenvector of A. Q.E.D.

The final theorem establishes necessary and sufficient conditions for the solution of the equation

$$A^{*}QA - Q = R \tag{8.13}$$

to be singular when the eigenvalues of A, $\lambda_i(A)$, are greater than one in absolute value and R is Hermitian, nonnegative definite. This solution can be written as

$$Q = \sum_{i=1}^{\infty} (A^{-i})^{*} R A^{-i} , \tag{8.14}$$

which implies that $Q > 0$ when $R > 0$. However, when R is positive semidefinite, Q may be either positive semidefinite or positive definite. The following theorem resolves this ambiguity.

Theorem 8.4:

If $|\lambda_i(A)| > 1$ and R is positive semidefinite, then Q is singular if and only if there exists an eigenvector of A which is null vector of R.

Proof:

$\leftarrow$ Let $\underline{x}$ be a vector such that $A\underline{x} = \lambda\underline{x}$ and $R\underline{x} = \underline{0}$. Then by pre- and postmultiplying (8.13) by $\underline{x}^{*}$ and $\underline{x}$, the equation

$$(\lambda^{*}\lambda - 1)\underline{x}^{*}Q\underline{x} = 0 \tag{8.15}$$

is obtained, and since $|\lambda| > 1$, (8.14) implies that $\underline{x}^{*}Q\underline{x} = 0$. Thus since Q is nonnegative definite, it is singular.

→ Let $\underline{n}$ be a null vector of Q and let $\underline{x} = A^{-1}\underline{n}$. Then (8.13) implies that

$$\underline{x}^*(Q + R)\underline{x} = \underline{n}^* Q\underline{n} = 0, \qquad (8.16)$$

and since both Q and R are nonnegative definite, $\underline{x}$ must be a null vector of both Q and R. Therefore, by repetition of this argument, it can be seen that every vector in the sequence $\{\underline{x}, A^{-1}\underline{x}, A^{-2}\underline{x}, \ldots\}$ is a null vector of both Q and R. These vectors define a cyclic subspace which, by the previous theorem, contains an eigenvector, $\underline{v}$, of A^{-1} which is a null vector of R. But $\underline{v}$ is also an eigenvector of A, which implies the desired conclusion.

IX. APPENDIX C: OBSERVABILITY AND CONTROLLABILITY THEOREMS

In this appendix, some tests for determining whether a given mode of a random process is completely driven and/or completely observable are stated and proved. These tests are considerably easier to apply than the basic definitions.

In the continuous-time case, the random process is described by the equations

$$\dot{\underline{x}} = A\underline{x} + \underline{h}(t) \tag{9.1}$$

$$\underline{y} = M\underline{x} \tag{9.2}$$

where $E[\underline{h}(t)\underline{h}'(u)] = H\delta(t-u)$. The eigenvalues of A are $\lambda_1, \ldots, \lambda_m$, and the Jordan decomposition of A is

$$AT = T\Lambda \tag{9.3}$$

where the matrices T and Λ can be partitioned as

$$T = [T_1, \ldots, T_m] \tag{9.4}$$

$$\Lambda = \mathrm{diag}(\Lambda_1, \ldots, \Lambda_m) \ . \tag{9.5}$$

By (9.3), the Jordan decomposition of A' is

$$A'(T^{-1})^* = (T^{-1})^*\Lambda^* \tag{9.6}$$

where $(T^{-1})^*$ can be partitioned as

$$(T^{-1})^* = [S_1, \ldots, S_m]. \tag{9.7}$$

The state space, $\mathcal{S}$, may be decomposed into the direct sums

$$\mathcal{S} = \tau_1 \oplus \cdots \oplus \tau_m$$
$$= \sigma_1 \oplus \cdots \oplus \sigma_m \tag{9.8}$$

where

τ_i = subspace spanned by the columns of T_i.

σ_i = subspace spanned by the columns of S_i.

Equations (9.4, 9.7) imply that $\tau_i \perp \sigma_j$ for all $i \neq j$. The ith mode, $\underline{\xi}_i(t)$, is defined to be completely driven if $\tau_i \subset \mathcal{R}(R_c)$, where

$$R_c = \int_{t_0}^{t_1} \Phi(t) H \Phi'(t) dt, \tag{9.9}$$

and it is defined to be completely observable if $\tau_i \cap \eta(P_c') = \underline{0}$, where

$$P_c = [M', A'M', \ldots, (A')^{n-1}M']. \tag{9.10}$$

Theorem 9.1:

The ith mode, $\underline{\xi}_i(t)$, of (9.1) is completely driven if and only if no eigenvector of A' corresponding to the eigenvalue λ_i^* is a null vector of H.

Proof:

The proof has two parts. In part (a) it is shown that $\underline{\xi}_i(t)$ is completely driven if and only if $\sigma_i \cap \eta(R_c) = \underline{0}$, i.e. σ_i and $\eta(R_c)$ have no nontrivial vectors in common, and in part (b) it is shown that σ_i has a nontrivial vector in $\eta(R_c)$ if and only if there is an eigenvector of A' in σ_i which is a null vector of H.

(a) $\leftarrow$ If $\sigma_i \cap \eta(R_c) = \underline{0}$, then $\eta(R_c) \subset \sigma_1 \oplus \cdots \oplus \sigma_{i-1} \oplus \sigma_{i+1} \oplus \cdots \oplus \sigma_m = \tau_i^{\perp}$. Therefore $\tau_i \subset \eta(R_c)^{\perp} = \mathcal{R}(R_c)$.

$\rightarrow$ If $\underline{\xi}_i(t)$ is completely driven, then $\tau_i \subset \mathcal{R}(R_c)$ and $\tau_i^{\perp} = \sigma_1 \oplus \cdots \oplus \sigma_{i-1} \oplus \sigma_{i+1} \oplus \cdots \oplus \sigma_m \supset \mathcal{R}(R_c)^{\perp} = \eta(R_c)$, which implies that $\sigma_i \cap \eta(R_c) = \underline{0}$.

(b) $\leftarrow$ Let $\underline{n}$ be an eigenvector in σ_i such that $H\underline{n} = \underline{0}$. Then

$$\begin{aligned} R_c\underline{n} &= \int_{t_0}^{t_1} \Phi(t)H\Phi'(t)\underline{n}dt \\ &= \int_{t_0}^{t_1} e^{\lambda_i^* t}\, \Phi(t)H\underline{v}dt \\ &= \underline{0}. \end{aligned} \tag{9.11}$$

Thus $\underline{n}$ is a nontrivial vector in $\eta(R_c)$. $\rightarrow$ Let $\underline{v}$ be a nontrivial vector in σ_i such that $R_c\underline{v} = \underline{0}$. Since the integrand in (9.9) is nonnegative definite, $\Phi'(t)\underline{v}$ must be a null vector of H for all values of t. But

$$\Phi'(t) = \sum_{j=0}^{m-1} \phi_j(t)(A')^j$$

where m is the degree of the minimal polynomial of A, and $\phi_j(t)$ are linearly independent functions of time.

This implies that each of the vectors $\underline{v}_j \triangleq (A')^j \underline{v}$ is a null vector of H. Since σ_i is invariant under A', these vectors are all elements of σ_i and they define a cyclic subspace of σ_i which by Theorem 8.3, contains an eigenvector of A'. This eigenvector is a linear combination of the vectors $\underline{v}_j$, which implies that it is also a null vector of H. Q.E.D.

<u>Theorem 9.2</u>:

The mode $\underline{\xi}_i(t)$ is completely observable if and only if no eigenvector of A corresponding to the eigenvalue λ_i is a null vector of M.

<u>Proof</u>:

The following proof shows that $\underline{\xi}_i(t)$ is not completely observable if and only if there is an eigenvector in τ_i which is a null vector of M. This therefore also proves the theorem as stated.

$\leftarrow$ Let $\underline{v}$ be an eigenvector in τ_i such that $M\underline{v} = \underline{0}$. Then $P_c{}'\underline{v} = \underline{0}$, so $\underline{v}$ is a nontrivial vector common to both τ_i and $\eta(P_c{}')$. The mode, therefore, is not completely observable.

$\rightarrow$ Suppose $\underline{\xi}_i(t)$ is not completely observable. Then there must be a nontrivial vector, $\underline{n}$, in τ_i such that $P_c{}'\underline{n} = \underline{0}$. This implies that each of the vectors $\underline{n}_j \triangleq A^j\underline{n}$ is a null vector of M. Since τ_i is invariant under A, these vectors are all elements of τ_i and they define a cyclic subspace of τ_i which, by Theorem 8.3, contains an eigenvector of A. This eigenvector is a linear combination of the vectors $\underline{v}_j$, which implies that it is also a null vector of M. Q.E.D.

In the discrete-time case, the random process is described by the equations

$$\underline{x}_{k+1} = \Phi \underline{x}_k + \underline{h}_k \tag{9.12}$$

$$\underline{y}_k = M\underline{x}_k \tag{9.13}$$

where $E[\underline{h}_i \underline{h}_j'] = H\delta_{ij}$. The Jordan decompositions of Φ and Φ' are

$$\Phi T = T\Lambda \tag{9.14}$$

$$\Phi'(T^{-1})^* = (T^{-1})^* \Lambda^*$$

respectively where the partitionings of T, Λ, and $(T^{-1})^*$ are given by (9.4, 9.5, and 9.7). The i<u>th</u> mode, $\underline{\xi}_i(t_k)$, is defined to be completely driven if $\tau_i \subset \mathcal{R}(R_d)$, where

$$R_d = \sum_{j=0}^{n-1} \Phi^j H (\Phi^j)' , \tag{9.16}$$

and it is defined to be completely observable if $\tau_i \cap \eta(P_d') = \underline{0}$, where

$$P_d = [M', \Phi' M', \ldots (\Phi')^{n-1} M'] . \tag{9.17}$$

<u>Theorem 9.3</u>:

The i<u>th</u> mode of (9.12) is completely driven if and only if no eigenvector of Φ' corresponding to the eigenvalue λ_i^* is a null vector of H.

Proof:

The proof is very similar to the proof of Theorem 9.1. Part (a) is identical except for the replacement of R_c by R_d, and therefore is not repeated. In part (b) it is shown that σ_i has a nontrivial vector in $\eta(R_d)$ if and only if there is an eigenvector of Φ' in σ_i which is a null vector of H.

(b) $\leftarrow$ Let $\underline{n}$ be an eigenvector in σ_i such that $H\underline{n} = \underline{0}$. Then

$$\begin{aligned} R_d\underline{n} &= \sum_{j=0}^{n-1} \Phi^j H(\Phi^j)' \underline{n} \\ &= \sum_{j=0}^{n-1} (\lambda_i^*)^j \Phi^j H\underline{n} \\ &= \underline{0}, \end{aligned} \tag{9.18}$$

and $\underline{n}$ is a nontrivial vector in $\eta(R_d)$. $\rightarrow$ Let $\underline{v}$ be a nontrivial vector in σ_i such that $R_d\underline{v} = \underline{0}$. Since each term in the summation (9.16) is nonnegative definite, each of the vectors $\underline{v}_j = (\Phi')^j\underline{v}$ must be a null vector of H, and these null vectors define a cyclic subspace of σ_i which contains an eigenvector of Φ' which is a null vector of H. Q.E.D.

Theorem 9.4:

The mode $\underline{\xi}_i(t_k)$ is completely observable if and only if no eigenvector of Φ corresponding to the eigenvalue λ_i is a null vector of M.

Proof:

Except for the replacement of P_c by P_d and A by Φ, the proof of this theorem is identical to the proof of Theorem 9.2.

The final theorem establishes formulas which describe the subspaces of the state space which are completely undriven and completely unobservable respectively.

Theorem 9.5:

(a) The basis vectors for $\eta(R_d)$ can be chosen such that

$$\Phi' N_d = N_d \Lambda_d^* \tag{9.19}$$

and

$$HN_d = 0 \tag{9.20}$$

where the columns of N_d are the chosen basis vectors.

(b) The basis vectors of $\eta(P_d')$ can be chosen such that

$$\Phi N_0 = N_0 \Lambda_0 \tag{9.21}$$

and

$$MN_0 = 0 \tag{9.22}$$

where the columns of N_0 are the chosen basis vectors.

Proof:

(a) Let $\underline{n}$ be any null vector of R_d. Then since each term in (9.16) is nonnegative definite, $(\Phi')^j \underline{n}$ is a null vector of H for j=0, 1, ... n-1. The Cayley-Hamilton theorem implies that $(\Phi')^n\underline{n}$ is also a null vector of H, so therefore $R_d\Phi'\underline{n} = \underline{0}$. Thus $\eta(R_d)$ is invariant under Φ', and its basis vectors satisfy an equation such as (9.19). If the resulting Λ_d is not in Jordan form, it can be put into Jordan form by means of a similarity transformation. Equation (9.20) results from the fact that any null vector of R_d is also a null vector of H.

(b) Let $\underline{n}$ be any null vector of P_d'. Then $\Phi^j\underline{n}$ is a null vector of M for j = 0,1,...,n-1. The Cayley-Hamilton theorem implies that $\Phi^n\underline{n}$ is also a null vector of M, so therefore $P_d\Phi\underline{n} = 0$. Thus $\eta(P_d)$ is invariant under Φ, and its basis vectors satisfy (9.21). Again Λ_0 can be put into Jordan form by means of a similarity transformation, and (9.22) results from the fact that any null vector of P_d is also a null vector of M.